

**Structural Variations on the Turn Unit of DNA-Binding  
Hairpin Py-Im Polyamides**

Thesis by

Michelle Elizabeth Farkas

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2010

(Defended September 4, 2009)

© 2010

Michelle Elizabeth Farkas

All Rights Reserved



*For my family*

## Acknowledgements

I would like to thank my advisor, Peter Dervan, for providing a wonderful environment in which to perform research, and for all of the insight he has offered me with regard to projects, academia, and life in general. I would also like to thank the members of my thesis committee, Dennis Dougherty, Judy Campbell, and Brian Stoltz, for all of their support (and letters of recommendation) over the last five years. Additionally, I would like to thank Jackie Barton, who I have had the privilege of being a teaching assistant for for several years of my graduate career; it was truly a pleasure.

The Dervan lab has been a wonderful place to spend five years of graduate school, and the people in it made all the difference. I have had the opportunity to work with David Chenoweth, Christian Dose, Claire Jacobs, Dan Harki, Ben Li, Katy Muzikar, Nick Nickols, and Sherry Tsai on a number of projects. My collaboration with Sherry was a very important and fruitful one – I will never forget Team Chl and all of the trips to Scripps during my candidacy. I have thoroughly enjoyed being roomies with John Phillips, and his predecessor, Eric Fechter, who taught me a lot when I first joined the lab. I would also like to thank Ray Doss and Adam Poulin-Kerstein, two other fifth-years during my first year, for their friendship and wisdom offered throughout my own graduate career. But perhaps most importantly, I have come to value the friendships of Justin Cohen, Mareike Goeritz, and Katy Muzikar – no matter what corners of the world we end up in, I hope that we will always stay in touch. I like to thank the other members of the lab as well – Mike Brochu, Dan Gubler, Carey Hsu, Michael Marques, Dave Montgomery, Julie Poposki, Jim Puckett, Jim Sanchez, Ryan Stafford, Anne Viger, and Fei Yang. I have truly enjoyed the time that I have spent with all of you.

I would like to acknowledge my collaborators, James Chou and Joel Gottesfeld, at The Scripps Research Institute for all of their help and hosting while working on joint projects in La Jolla, Pasadena, and over the phone/internet. I would also like to thank the staff whom I have had the opportunity to interact with at Caltech. Mona Shahgholi at the

Mass Spec facility has been extremely helpful, along with Joe Drew, Steve Gould, Lillian Kremar, Mo Renta, and many others. I would like to offer a special note of gratitude to Lynne Martinez who has been incredible in her administrative role in the group.

There are many other people outside of the chemistry department whom I would like to thank for helping me get here. Thanks to all of my roommates, Snatch teammates, and other friends throughout the years at Caltech – it's been really nice to have somewhere to go outside of lab. Thanks to all of my high school and Wellesley friends who've watched me go through this and been as supportive as they have, despite my falling off of the face of the earth every so often. Thanks to Pam Sontz for not only being a great friend, but also reading this thesis. And thanks to Carolyn Brinkworth for both being and doing so much more than that. I would also like to thank my former mentor Dr. Sandor Karady at Merck who allowed me to do research as a high school student; he set me on a course that brought me to where I am today. Dr. Kap-Sun Yeung was a wonderful person to work for during my years at Bristol-Myers Squibb.

Lastly, I would like to thank my parents and family for all of their support leading up to and during my graduate career. There have been a lot of phone calls where I haven't had much to say, but I would like to share this achievement with all of you. I am incredibly grateful for all that you have done for me throughout the years and love you very much.

## Abstract

Modulation of gene expression by small molecules is a challenge in the field of chemical biology. Hairpin pyrrole-imidazole polyamides are a class of programmable small molecules that bind to the minor groove of DNA, and have been shown to inhibit gene expression by interfering with transcription factor-DNA interfaces. When considering the biological implications of these molecules, improvement of sequence-specificities and binding affinities is of great interest. The work described herein focuses on modifications to the turn sub-unit of hairpin polyamides, and subsequent effects on the biophysical and biological characteristics of these molecules. The substitution of a  $\gamma$ -2,4-diaminobutyric acid with an  $\alpha$ -2,4-diaminobutyric acid hairpin turn moiety resulted in greater selectivity, and diminished reactivity for polyamide-alkylator conjugates. These molecules have been utilized in generating site-specific damage in histone H4 genes in cancer cells. Employment of 3,4-diaminobutyric acid in the turn unit has resulted in increased DNA affinities for polyamides targeting particular sequences. These molecules are also promising in their abilities to tolerate modifications that improve cellular uptake but would otherwise severely diminish binding.

## Table of Contents

	Page
Acknowledgements.....	iv
Abstract.....	vi
Table of Contents.....	vii
List of Figures and Tables.....	viii
 Chapter I. Introduction.....	 1
 Part II. Py-Im Polyamides linked by 2,4-diaminobutyric acid in the hairpin turn unit.....	 22
Chapter IIA. Unanticipated differences between $\alpha$ - and $\gamma$ - diaminobutyric acid-linked hairpin polyamide-alkylator conjugates.....	 23
Chapter IIB. Characterization of $\alpha$ -diaminobutyric acid-linked hairpin polyamides.....	 49
Chapter IIC. Small molecules targeting histone H4 as potential therapeutics for chronic myelogenous leukemia.....	 79
 Part III. Substituent effects in the $\gamma$ turn unit of hairpin Py-Im Polyamides.....	 111
Chapter IIIA. Next generation hairpin polyamides with ( <i>R</i> )-3,4-diaminobutyric acid turn unit.....	 112
Chapter IIIB. DNA sequence selectivity of hairpin polyamide turn units.....	143
 Appendix. Exploration of methods for genome-wide analysis of polyamide binding.....	 172

## List of Figures and Tables

Chapter I	Page
Figure I.1    The structure of DNA.....	3
Figure I.2    Structures of five protein-DNA complexes.....	4
Figure I.3    Chemical structures of DNA-binding natural products.....	5
Figure I.4    Crystal structures of distamycin bound to DNA.....	6
Figure I.5    Recognition of the DNA minor groove by polyamides.....	8
Figure I.6    Crystal structures of polyamides bound to DNA.....	9
Figure I.7    Binding model for the polyamide CtPyPyIm-( <i>R</i> ) <sup>RHN</sup> γ-PyImPyPy-β-Dp targeted to the sequence 5'-TATACGT-3'.....	11
Table I.1    Library of imidazole-capped polyamides.....	13
Figure I.8    Polyamides as regulators of gene expression in cell culture.....	14
Figure I.9    Examples of various polyamide-alkylator conjugates.....	16
 <b>Chapter IIA</b>	
Figure IIA.1    Polyamide structures and syntheses.....	28
Figure IIA.2    Plasmid design and DNA binding properties of parent polyamides.....	30
Figure IIA.3    Illustration of parent polyamides <b>2R</b> and <b>2S</b> binding DNA.....	31
Table IIA.1    Binding affinities for parent polyamides.....	31
Figure IIA.4    Models of the binding of <b>1R</b> and <b>1S</b> to the match site of pMFST2.....	32
Figure IIA.5    Alkylation specificities of the polyamide conjugates.....	33
Figure IIA.6    Time-dependance of alkylation for polyamide conjugates.....	35
Figure IIA.7    Effects of parent polyamides on cultured SW620 cells.....	37

Figure IIA.8	Effects of polyamide conjugates on cultured cells.....	38
Figure IIA.9	Fluorescence-activated cell-sorting analysis of the effects of polyamide conjugates on SW620 cells.....	40
Figure IIA.10	Effects of polyamide conjugates on BALB/c mice.....	41

## Chapter IIB

Figure IIB.1	Structures of polyamides <b>1-8</b> and their <b>Chl</b> conjugates.....	53
Figure IIB.2	Illustration of the inserts from plasmids pMFST2, pMFST, and pMFST3.....	55
Figure IIB.3	Quantitative DNase I footprinting titration experiments for Polyamides <b>1-8</b> .....	57
Table IIB.1	Binding affinities for parents polyamides on plasmids pMFST2, pMFST, and pMFST3.....	58
Figure IIB.4	Thermal cleavage assay experiments with polyamide-chlorambucil conjugates.....	60
Figure IIB.5	Illustration of the binding modes for polyamides <b>3</b> and <b>7</b> .....	61
Figure IIB.6	Structures of <b>EDTA-3</b> , and the glycine-linked polyamide <b>9</b> and its EDTA conjugate <b>EDTA-9</b> .....	63
Figure IIB.7	Experiments involving polyamides <b>3</b> , <b>9</b> , <b>EDTA-3</b> , and <b>EDTA-9</b> on plasmid pMFST5.....	64
Table IIB.2	Binding affinities for parent polyamides <b>3</b> and <b>9</b> on plasmid pMFST5.....	65
Figure IIB.8	Affinity cleavage experiments for <b>EDTA-7</b> and <b>EDTA-8</b> .....	67
Figure IIB.9	Affinity cleavage experiments for <b>EDTA-10</b> on plasmid pMFST4.....	68
Figure IIB.10	Structures of polyamide-fluorescein conjugates.....	70
Table IIB.3	Cellular uptake properties of polyamide-fluorescein conjugates.....	71

## Chapter IIC

Figure IIC.1	DNA sequence of the coding region of the histone H4C gene and chemical structures for <b>1R-Chl</b> and <b>6R-Chl</b> .....	83
Table IIC.1	Analysis of a polyamide library.....	84
Figure IIC.2	Polyamide analysis on pMFST6.....	85
Figure IIC.3	Alkylation of the H4c gene <i>in vitro</i> .....	87
Figure IIC.4	Effects of polyamides in cell culture.....	88
Figure IIC.5	Analysis of inactive polyamide in K562 cells.....	89
Figure IIC.6	Alkylation of the H4c gene in K562 cells.....	91
Figure IIC.7	Effect of polyamides on histone H4c and H4k/j transcript expression.....	92
Figure IIC.8	Real time qRT-PCR of genes.....	94
Figure IIC.9	Murine K562 xenograft studies.....	95
Figure IIC.10	Pharmacokinetic parameters and biodistribution of <b>1R-Chl</b> .....	97
Table IIC.2	Primers used for real-time qRT-PCR.....	104

## Chapter IIIA

Figure IIIA.1	Schematic representation of hairpin polyamides with different turn units.....	115
Figure IIIA.2	Chemical structures for polyamides <b>1-16</b> .....	116
Figure IIIA.3	DNase I footprinting experiments for polyamides <b>2-4</b> .....	117
Figure IIIA.4	DNase I footprinting experiments for polyamides <b>6-8</b> .....	118
Figure IIIA.5	DNase I footprinting experiments for polyamides <b>10-12</b> .....	119
Figure IIIA.6	DNase I footprinting experiments for polyamides <b>14-16</b> .....	120
Table IIIA.1	Equilibrium association constants.....	121



Figure IIIA.7	Normalized UV denaturation profiles of DNA duplex with and without polyamides.....	122
Table IIIA.2	Melting temperatures of DNA/polyamide complexes.....	123
Figure IIIA.8	Illustrative models of different turn conformations.....	125
Table IIIA.3	Melting temperatures of DNA/polyamide complexes with C•G and G•C base pairs at the turn position.....	127
Figure IIIA.9	Chemical structures of acetylated polyamides.....	129
Table IIIA.4	Melting temperatures for DNA complexes with acetylated hairpins.....	129
Figure IIIA.10	Schematic representation of the androgen receptor transcription complex.....	130
Figure IIIA.11	Structures of polyamides <b>20-23</b> , and their effects on gene expression in cell culture.....	131
Table IIIA.5	Melting temperatures of polyamides <b>20-23</b> .....	132

## Chapter IIIB

Figure IIIB.1	Schematic diagram of a six-ring hairpin polyamide targeting DNA.....	146
Figure IIIB.2	Structures of polyamides targeting 5'-WWGGWW-3'.....	148
Figure IIIB.3	Structures of polyamides <b>6-11</b> .....	148
Table IIIB.1	Melting temperatures of DNA/polyamide complexes targeting 5'-WWGGWW-3'.....	149
Table IIIB.2	Melting temperatures of DNA/polyamide complexes targeting 5'-WWGWW-3' and 5'-WWGGCW-3'.....	150
Figure IIIB.4	Illustration of the designed inserts for plasmids pCDMF6 and pCDMF4.....	151
Figure IIIB.5	Quantitative DNase I footprinting experiments for polyamides targeting 5'-WWGGWW-3'.....	152
Table IIIB.3	Binding affinities for polyamides targeting 5'-WWGGWW-3'.....	153

Table IIIB.4	Relative binding affinities for polyamides targeting 5'-WWGGWW-3' .....	153
Figure IIIB.6	Quantitative DNase I footprinting experiments for polyamides targeting 5'-WWGWW-3' .....	154
Table IIIB.5	Binding affinities for polyamides targeting 5'-WWGWW-3' .....	155
Table IIIB.6	Relative binding affinities for polyamides targeting 5'-WWGWW-3' .....	155
Figure IIIB.7	Structures of polyamides containing fluoro- and hydroxyl-substituted hairpin turns.....	156
Table IIIB.7	Melting temperatures of DNA/polyamide complexes for polyamides containing fluoro- and hydroxyl-substituted hairpin turns.....	156
Figure IIIB.8	Quantitative DNase I footprinting experiments for polyamides containing fluoro- and hydroxyl-substituted hairpin turns.....	157
Table IIIB.8	Binding affinities for polyamides containing fluoro- and hydroxyl-substituted hairpin turns.....	158
Table IIIB.9	Relative binding affinities for polyamides containing fluoro- and hydroxyl-substituted hairpin turns.....	158
Figure IIIB.9	Synthesis of fluorine-substituted hairpin turns.....	161
Figure IIIB.10	Synthesis of hydroxyl-substituted hairpin turns.....	161

## Appendix

Figure A.1	Polyamides targeting the androgen response element.....	175
Figure A.2	Illustration of chromatin immunoprecipitation versus pull-down with polyamide-biotin conjugates.....	177
Figure A.3	Structures of polyamides linked to biotin at the turn subunit.....	178
Figure A.4	Effects of polyamides <b>1-3</b> on the expression of androgen responsive genes.....	179
Table A.1	Melting temperatures of DNA/polyamide complexes.....	180

Figure A.5	Structures of polyamides linked to biotin at the tail subunit.....	181
Table A.2	Melting temperatures of DNA/polyamide complexes.....	182
Figure A.6	Effects of polyamides <b>4-9</b> on PSA expression.....	183
Figure A.7	Electrophoretic mobility shift assay with polyamides <b>8</b> and <b>9</b> .....	183
Figure A.8	Structure of polyamide <b>10</b> .....	184
Figure A.9	Relative amplification levels of PSA ARE III and FKBP5 following polyamide incubation with sheared chromatin.....	185
Figure A.10	Relative amplification levels of PSA ARE I, PSA ARE III, and FKBP5 following polyamide treatment in LNCaP cells.....	186
Figure A.11	Relative amplification levels of loci containing 5'-WGWWCW-3' repeats.....	186

## **Chapter I**

### **Introduction**

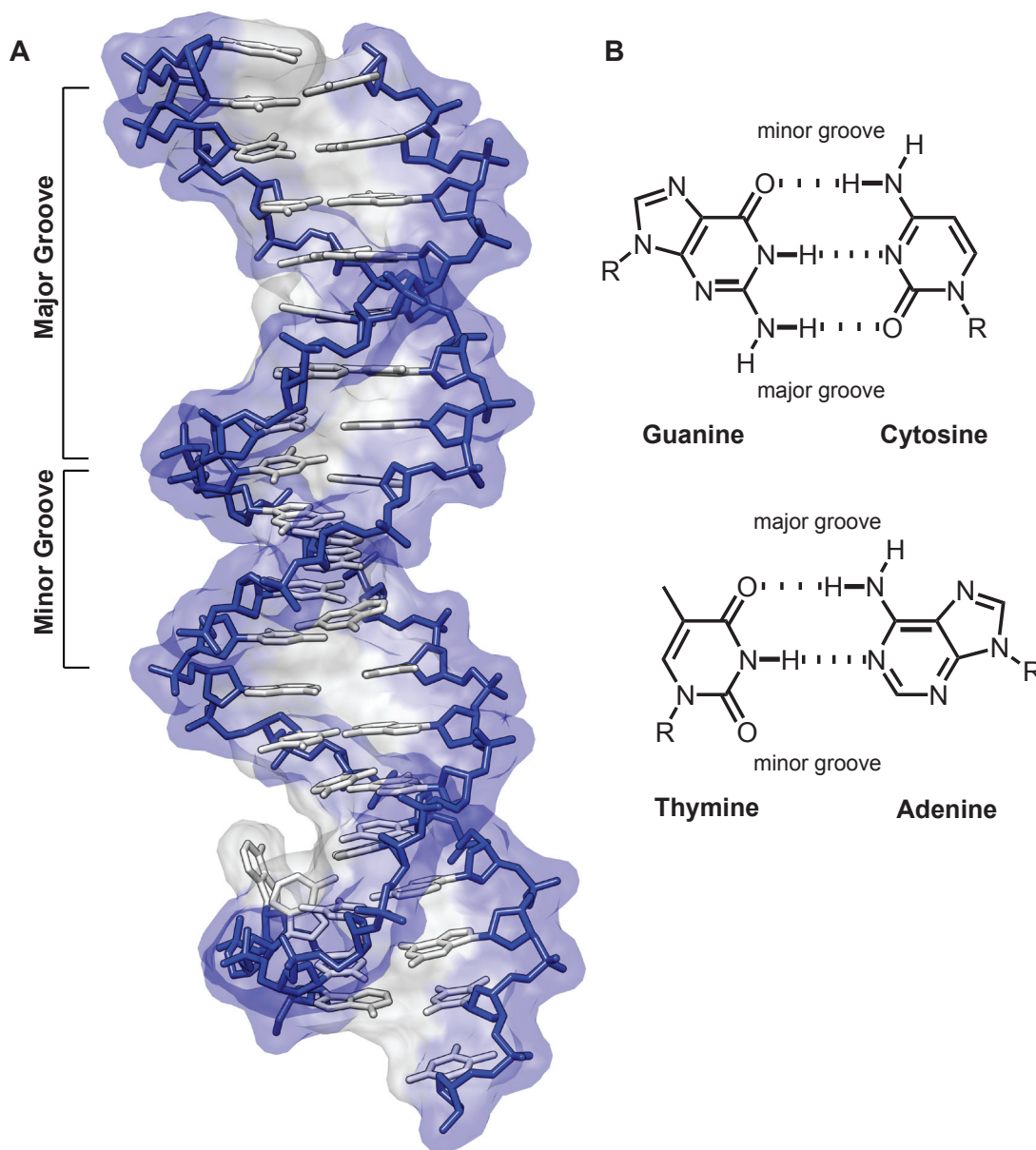
## **Deoxyribonucleic acid (DNA)**

The macromolecule DNA encodes the genetic information required for life. The human genome contains approximately three billion base pairs, which encode 20,000 to 25,000 protein-coding genes.<sup>1</sup> Sequences of DNA can not only code for proteins, but are also associated with regulatory regions controlling transcription. In the cell, gene expression is controlled by DNA-binding proteins and the interactions between proteins.<sup>2,3</sup>

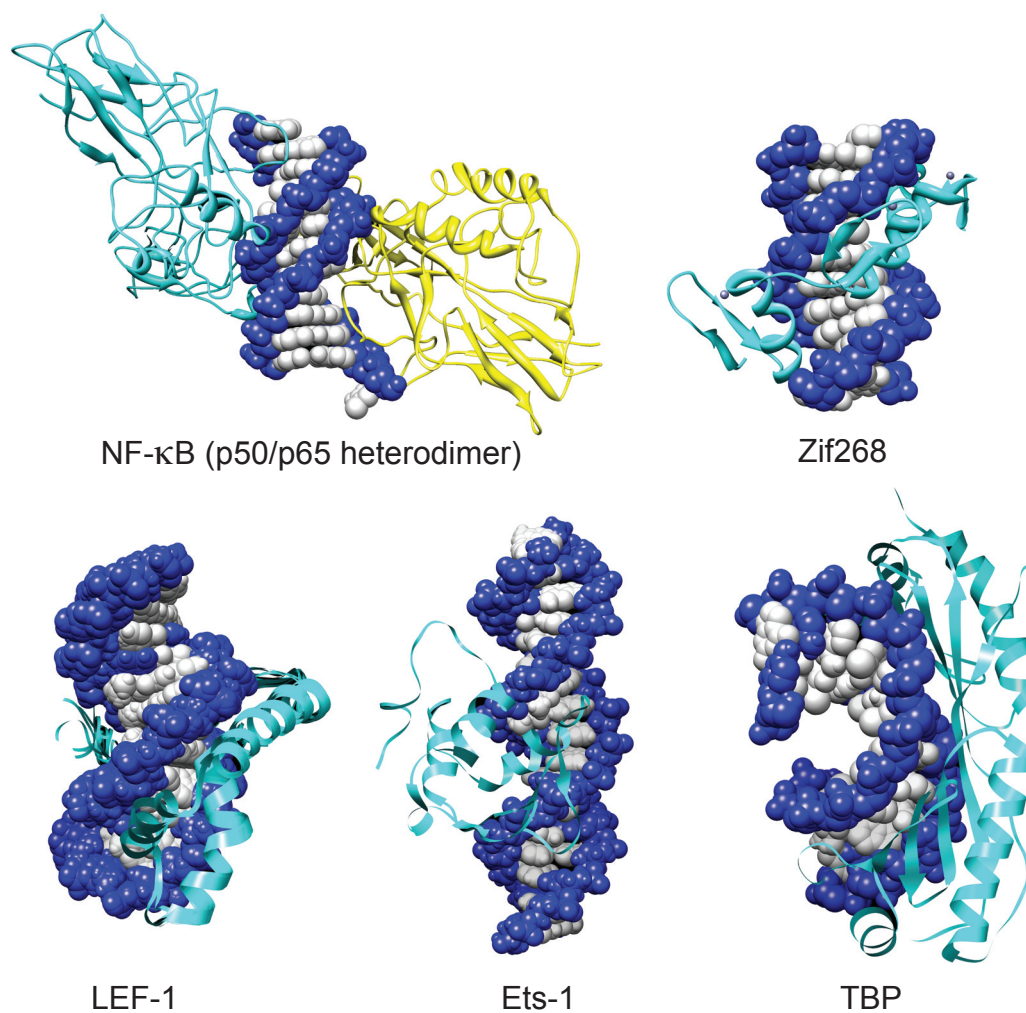
DNA consists of two anti-parallel polydeoxyribonucleotide strands wrapped together in a double helix. Between the two strands, hydrogen bonding of the Watson-Crick heterocyclic base pairs occurs, with adenine (A) pairing with thymine (T), and guanine (G) pairing with cytosine (C). B-form DNA is a right-handed helix containing ten base pairs per turn.<sup>4,5</sup> The helix forms two grooves: the major groove, which is wide and shallow, and the minor groove, which is narrow and deep (Figure I.1). The edges of the DNA base pairs present unique molecular surfaces at the major and minor groove floors, providing for the sequence-specific recognition of DNA by transcription factors and other molecules.

## **DNA-binding small molecules**

The transcription of DNA into mRNA results in the synthesis of proteins, or translation. Many proteins function as transcription factors, regulating gene expression. These proteins must recognize and bind particular sequences of DNA, whether in the minor, major, or a combination of the two grooves (Figure I.2). NF- $\kappa$ B (p50/p65 heterodimer),<sup>6</sup> Zif268,<sup>7</sup> and Ets-1<sup>8</sup> recognize the major groove, while TATA-Binding Protein (TBP)<sup>9</sup> and LEF-1<sup>10</sup> recognize the minor groove of DNA.

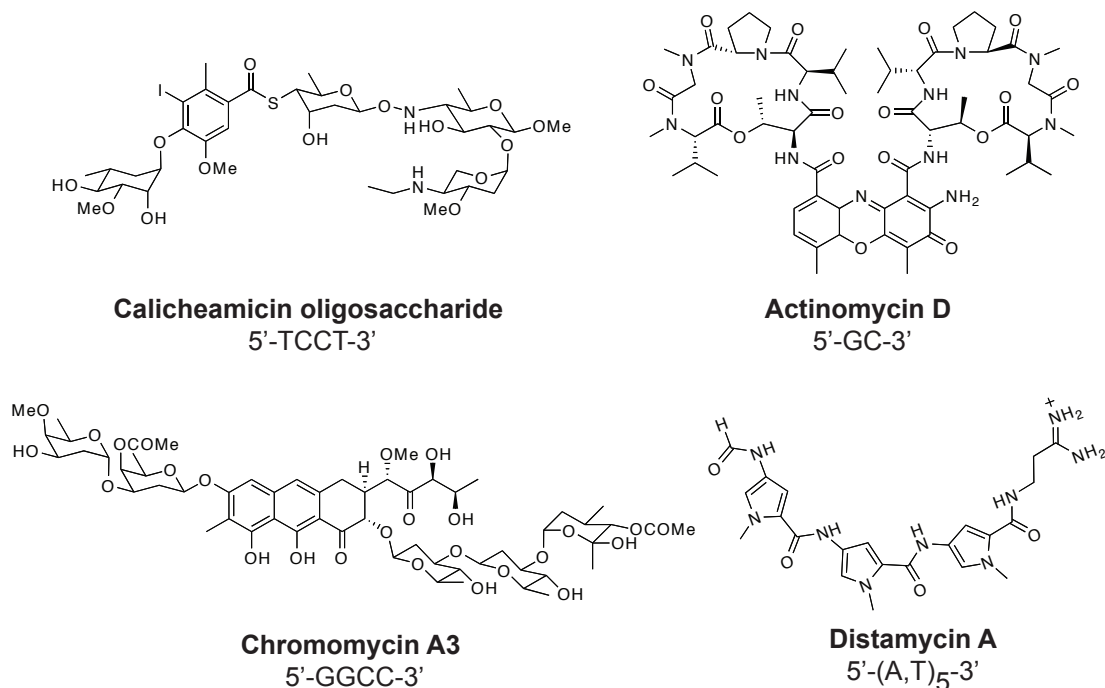


**Figure I.1.** The structure of DNA. **(A)** X-ray crystal structure of B-form DNA. The phosphodiester-linked deoxyribose backbone is shown in blue, and the Watson-Crick base pairs are shown in gray (PDB 1YSA).<sup>3</sup> **(B)** Chemical structures of hydrogen-bonded base pairs. Cytosine (C) is bonded to guanine (G) (top), and thymine (T) is bonded to adenine (A) (bottom). 'R' = the sugar phosphosphate backbone of DNA. Dashed lines indicate hydrogen bonds. The major and minor grooves are indicated in both representations.



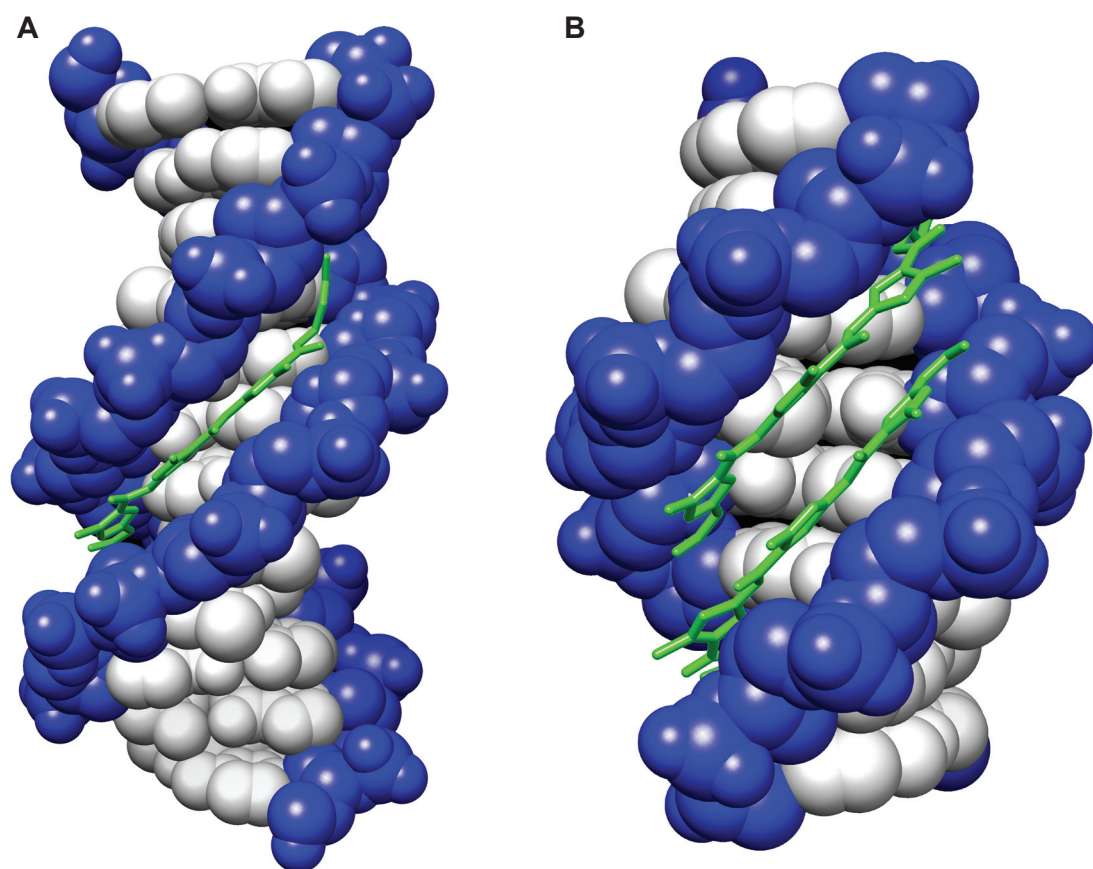
**Figure I.2.** Structures of five protein-DNA complexes: NF- $\kappa$ B (PDB 1VKX),<sup>6</sup> Zif268 (PDB 1ZAA),<sup>7</sup> LEF-1 (PDB 2LEF),<sup>10</sup> Ets-1 (PDB 2STW),<sup>8</sup> and TBP (PDB 1TGH).<sup>9</sup> All protein structures are determined by X-ray crystallography with the exception of Ets-1, which was determined by NMR.

In addition to proteins, a number of small molecule natural products are capable of binding to specific sequences of DNA. Calicheamicin,<sup>11</sup> actinomycin D,<sup>12</sup> chromomycin<sup>13</sup> and distamycin A<sup>14</sup> have been shown to bind DNA with high affinity and modest sequence specificity within the minor groove (Figure I.3). Chromomycin targets the sequence 5'-GGCC-3' and binds in the minor groove in a 2:1 stoichiometry. Actinomycin D intercalates DNA at 5'-GC-3' sequences in a 1:1 stoichiometry. Calicheamicin oligosaccharide binds the minor groove as a monomer as well, and targets 5'-TCCT-3'. Distamycin A is an A,T-binding oligopeptide whose structure contains three *N*-methylpyrrole (Py) carboxamide units. It was revealed via x-ray and NMR structural studies that two of these crescent-shaped molecules can bind DNA as a 1:1 or 2:1 complex in an anti-parallel orientation (Figure I.4). This results in expansion of the minor groove relative to the 1:1 ligand-DNA complex.<sup>15-17</sup>



**Figure I.3.** Chemical structures of DNA-binding natural products. Structures of calicheamicin oligosaccharide, actinomycin D, chromomycin A3, and distamycin A are shown with targeted DNA sequences.

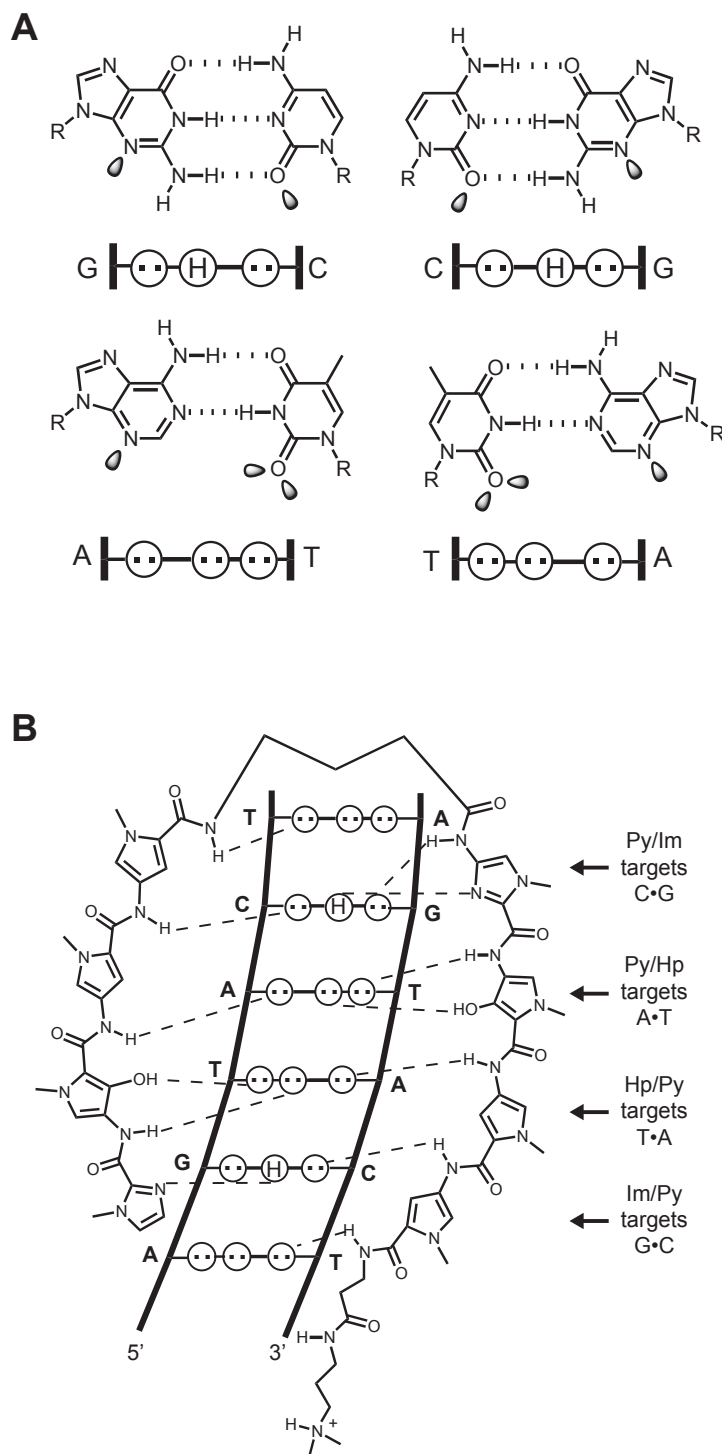




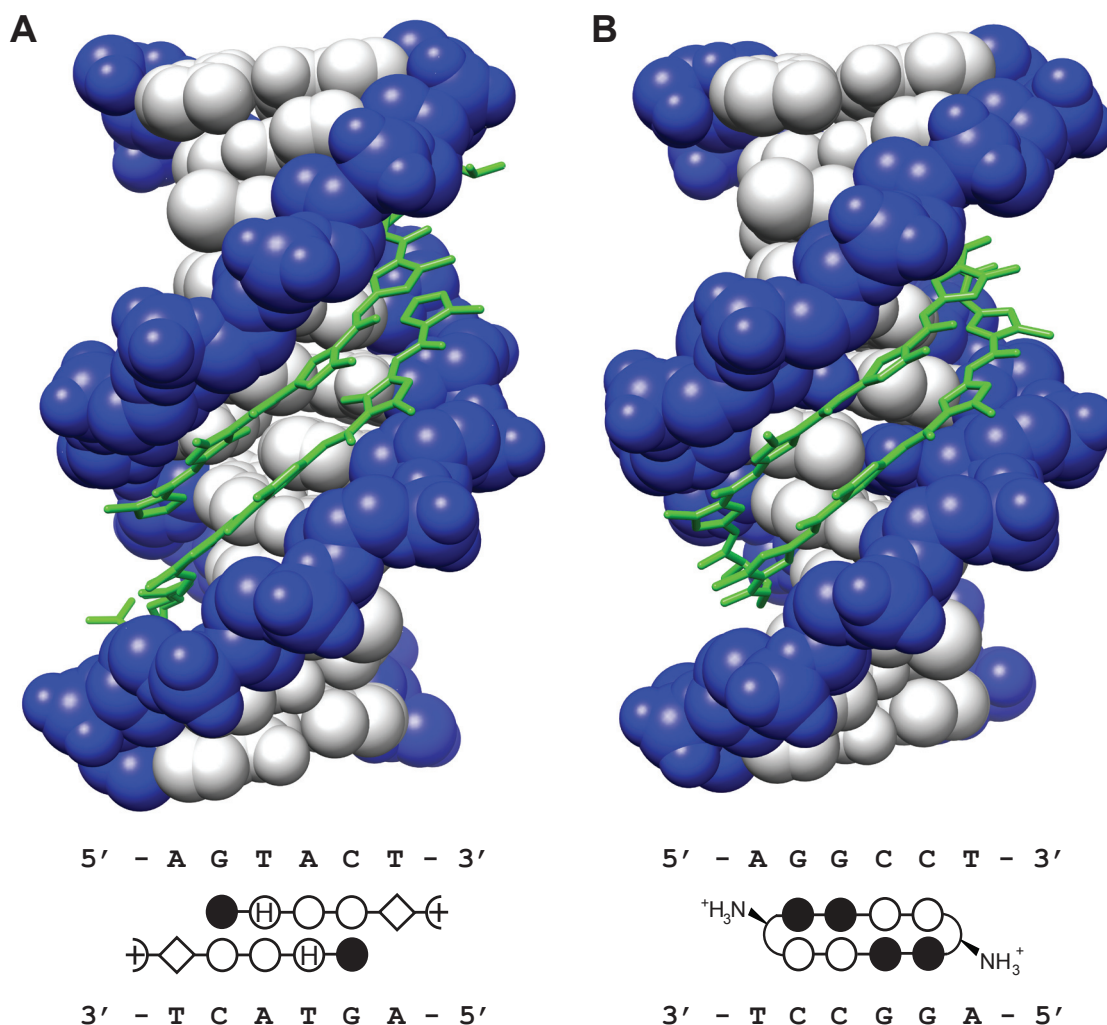
**Figure I.4.** Distamycin bound to DNA. X-ray crystal structures of distamycin bound to DNA in (A) 1:1 (PDB 2DND)<sup>15</sup> and (B) 2:1 (right, PDB 378D)<sup>17</sup> conformations are shown.

### Minor groove recognition by pyrrole-imidazole polyamides

Pyrrole-imidazole polyamides bind in the minor groove of DNA and are able to discern the four Watson-Crick base pairs based upon the chemical features presented.<sup>18,19</sup> The exocyclic amine of guanine presents a steric clash to *N*-methylpyrrole (Py), however the nitrogen of *N*-methylimidazole (Im) is capable of hydrogen bonding to it. Subsequently, a set of “pairing rules” have been developed, where a Py/Im pair targets C•G, the reverse (Im/Py) targets G•C, and Py/Py targets A•T or T•A.<sup>20</sup> The degeneracy between A and T is successfully broken by usage of *N*-methylhydroxypyrrole (Hp), on account of its hydroxyl group’s ability to form hydrogen bonds with the O2 of thymine (Figure I.5).<sup>21,22</sup> Alternatively, a terminal 3-chlorothiophene (Ct) also shows a three-fold selectivity for T•A over A•T when paired opposite Py.<sup>23</sup> The replacement of Py with  $\beta$ -alanine ( $\beta$ ) can enhance polyamide binding affinity by allowing the polyamide to reset its curvature with that of the DNA minor groove.<sup>24,25</sup>  $\beta$ /Py and  $\beta$ / $\beta$  pairings both recognize T•A and A•T base pairs. A  $\beta$  in the polyamide tail position and the 3-(dimethylamino)propylamine (Dp) tail itself also both code for T•A and A•T. Structural verification of the pairing rules has been provided by X-ray crystal structures of polyamides bound to DNA (Figure I.6).<sup>20,22,26</sup>



**Figure I.5.** Recognition of the DNA minor groove by polyamides. (A) Minor groove hydrogen bonding patterns of Watson-Crick base pairs. Circles with dots represent lone pairs of N(3) of purines and O(2) of pyrimidines, and circles containing an H represent the 2-amino group of guanine. 'R' represents the sugar-phosphate backbone of DNA. Shaded orbitals represent electron pairs projecting into the minor groove. (B) Model of the polyamide ImHpPyPy- $\gamma$ -ImHpPyPy- $\beta$ -Dp binding its target sequence, 5'-AGTACT-3'. Putative hydrogen bonds are shown as dashed lines.

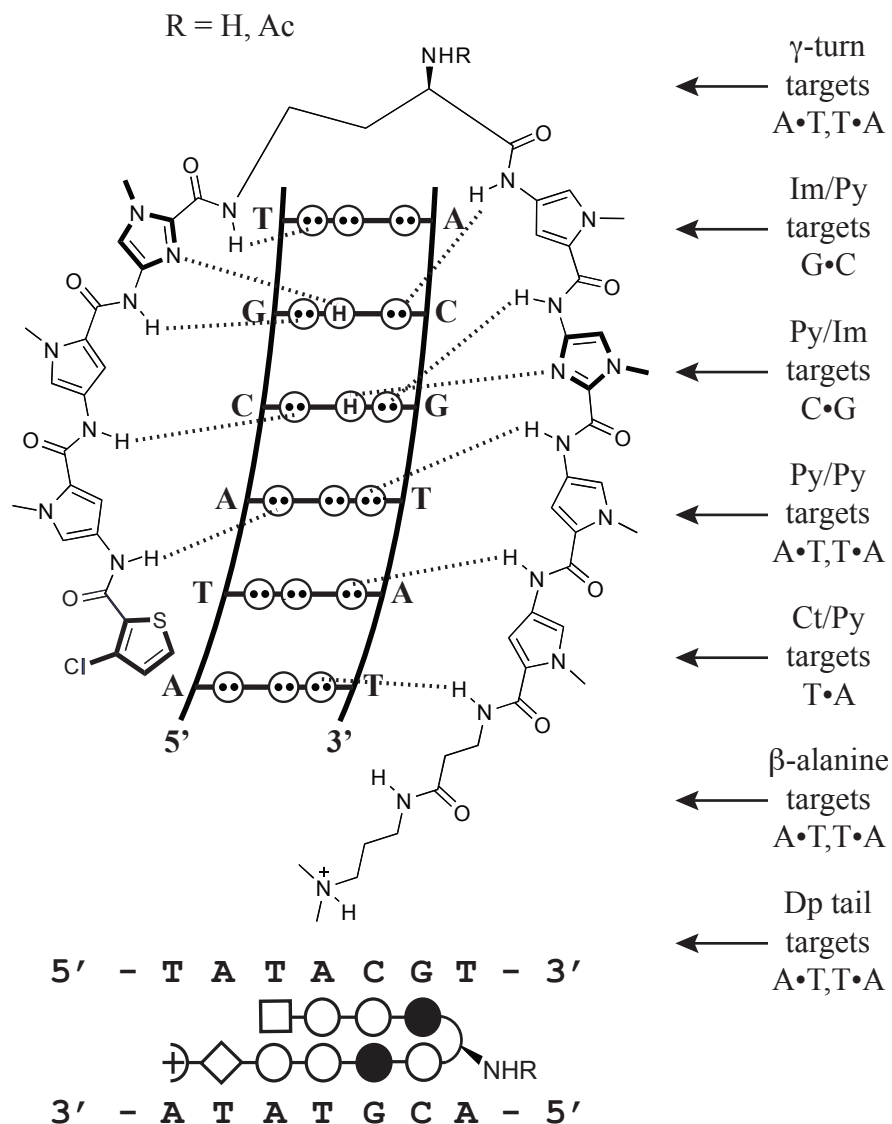


**Figure I.6.** Crystal structures of polyamides bound to DNA. (A) Structure of the 2:1 complex between ImHpPyPy-β-Dp and the sequence 5'-CCAGTACTGG-3' (PDB 407D).<sup>22</sup> (B) Structure of the cyclic polyamide ImImPyPy-(*R*) α-amine-γ-ImImPyPy-(*R*) α-amine-γ bound to 5'-CCAGGCCTGG-3' (PDB 3I5L).<sup>26</sup> Polyamides are shown below in ball-and-stick models, where a black circle represents Im, an open circle represents Py, a circle containing 'H' represents Hp, a diamond represents β-alanine, a half-circle with a plus represents Dp, and a half-circle represents γ-turn.

### Hairpin motifs in Py-Im polyamides

By binding as side-by-side, anti-parallel dimers in the minor groove, the heterocyclic rings of two polyamide strands can recognize DNA via the pairing rules described above.<sup>27</sup> Covalent linkage of the carboxy terminus (C) of one strand to the amino terminus (N) of another via an alkyl turn moiety results in a 'hairpin' polyamide structure (Figure I.7). Early studies explored a variety of linker lengths and revealed that the use of a  $\gamma$ -aminobutyric acid ( $\gamma$ -turn) is optimal, and resulted in increased affinities ( $\sim 100$ -fold) and specificities compared to unlinked subunits.<sup>28</sup> The  $\gamma$ -turn also enforces the register of the heterocyclic ring pairings, preventing slipped dimer binding modes.<sup>29</sup>

The hairpin moiety also contributes to the polyamide's recognition of particular DNA sequences. The  $\gamma$ -turn prefers to bind at T•A and A•T base pairs over G•C and C•G.<sup>30</sup> Usage of a chiral (*R*)-2,4-diaminobutyric acid in the hairpin further improves binding affinity ( $\sim 10$ -fold over the  $\gamma$ -turn), and retains specificity for T•A and A•T.<sup>31</sup> Hairpin polyamides containing the  $\gamma$ -turn typically prefer to bind in the 'forward' direction N $\rightarrow$ C with respect to the 5' $\rightarrow$ 3' direction of the DNA strand,<sup>32</sup> however in some instances the 'reverse' (C $\rightarrow$ N alignment with respect to the 5' $\rightarrow$ 3' direction of the DNA strand) orientation is preferred. The chiral turn adds stereochemical control, enforcing polyamide binding in the forward orientation.<sup>31</sup>



**Figure I.7.** Binding model for the polyamide CtPyPyIm-(R)<sup>RHN</sup>γ-PyImPyPy-β-Dp targeted to the sequence 5'-TATACGT-3'. The heterocycles Im and Ct are shown in bold. A ball-and-stick model is shown below, whereby a square represents Ct, a black circle represents Im, an open circle represents Py, a diamond represents β-alanine, a half-circle with a plus represents Dp, and a half-circle represents γ-turn.

### **Polyamide regulation of gene expression in cell culture**

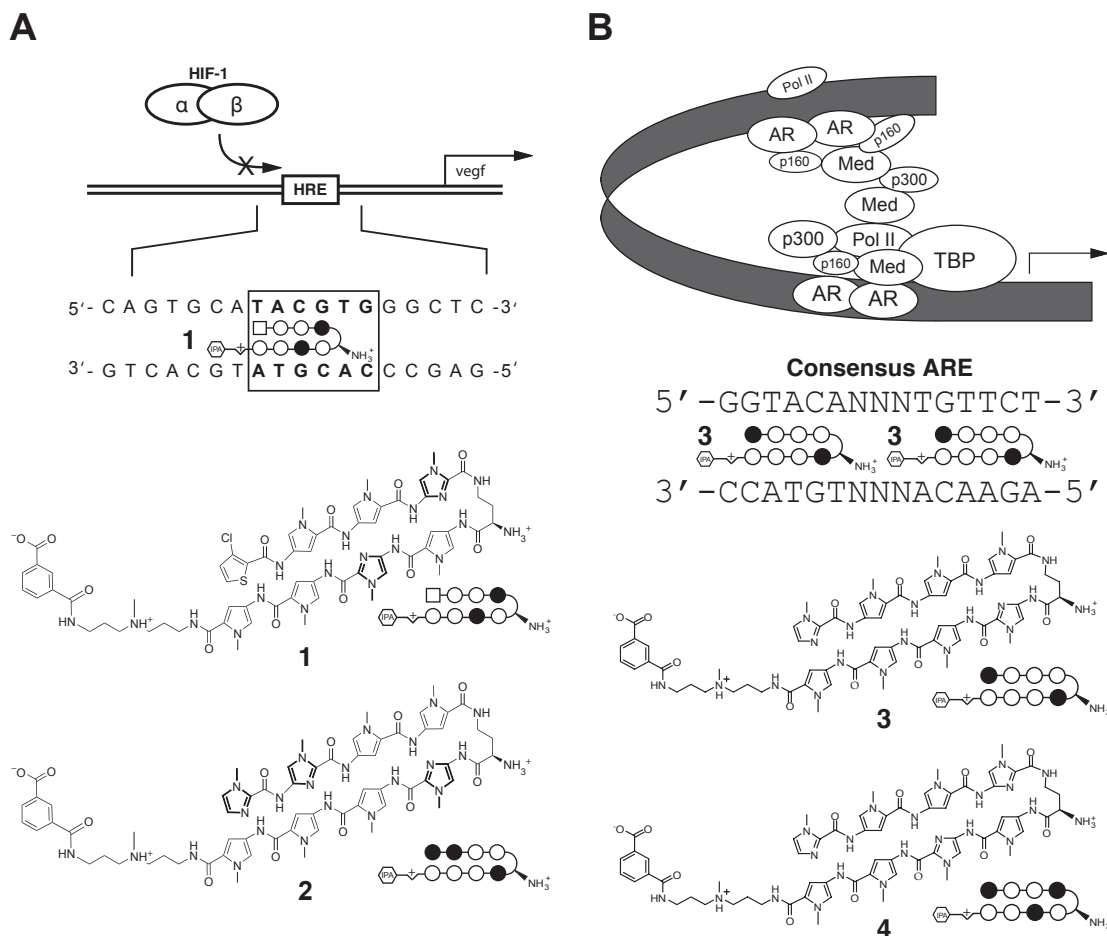
Eight-ring hairpin polyamides have been shown to bind a wide repertoire of DNA match sequences with subnanomolar affinities, similar to those of DNA-binding proteins (Table I.1).<sup>33-35</sup> By using Py-Im polyamides to displace or prevent the binding of transcription factors to their respective promoter DNA sequences, the expression of particular genes can be modulated. Cellular permeability is of vital importance to the successful design of gene regulation studies. Confocal microscopy has been used to determine the nuclear uptake profiles of polyamide-fluorophore conjugates in a variety of cell lines.<sup>36-38</sup> The presence of an isophthalic acid (IPA) moiety in the tail region has been shown to yield high affinity conjugates with improved nuclear permeability.<sup>39</sup>

Polyamides have been used to successfully affect gene expression in cell culture. A polyamide designed to bind to the hypoxia response element (HRE) was shown to disrupt the binding of hypoxia-inducible factor (HIF) to the HRE, resulting in decreased transcription of vascular endothelial growth factor (VEGF) in cultured HeLa cells (Figure I.8A).<sup>40,41</sup> Another polyamide, designed to target the androgen response element (ARE), has been shown to down-regulate prostate-specific antigen (PSA) and other androgen responsive genes in prostate cancer cells (Figure 1.8B).<sup>42</sup>

	Sequence (5'→3')	Polyamide	$K_a$ (M <sup>-1</sup> )	Context
1	WWGWWWW		$3 \times 10^9$	5' - TAGTATT - 3'
2	WWGGWWW		$5 \times 10^8$	5' - CTGGTTA - 3'
3	WWGWWGW		$4 \times 10^9$	5' - TAGTGAA - 3'
4	WWGWWGW		$9 \times 10^9$	5' - TAGTAGT - 3'
5	WWGWWCW		$3 \times 10^{10}$	5' - TAGTACT - 3'
6	WWGWCWW		$2 \times 10^9$	5' - GAGTCTA - 3'
7	WWGCWWW		$5 \times 10^9$	5' - ATGCAAA - 3'
8	WWGGGWW		$3 \times 10^8$	5' - AAGGGAA - 3'
9	WWGGWGW		$1 \times 10^{10}$	5' - TAGGTGT - 3'
10	WWGGWCW		$1 \times 10^{10}$	5' - ATGGTCA - 3'
11	WWGGCWW		$4 \times 10^8$	5' - AAGGCAT - 3'
12	WWGWGGW		$4 \times 10^{10}$	5' - TAGTGGT - 3'
13	WWGWGCW		$2 \times 10^9$	5' - ATGAGCT - 3'
14	WWGCGWW		$2 \times 10^9$	5' - ATGCGTA - 3'
15	WWGCWGW		$2 \times 10^9$	5' - TAGCAGT - 3'
16	WWGCWCW		$9 \times 10^9$	5' - ATGCTCA - 3'
17	WWGWCGW		$1 \times 10^{10}$	5' - ATGACGT - 3'
18	WWGWCCW		$2 \times 10^9$	5' - TAGACCA - 3'
19	WWGCCWW		$7 \times 10^8$	5' - ATGCCTA - 3'
20	WWGGGGW		$2 \times 10^8$	5' - GAGGGGT - 3'
21	WWGCGGW		$9 \times 10^8$	5' - ATGCGGT - 3'
22	WWGGCGW		$2 \times 10^8$	5' - CAGGCGT - 3'
23	WWGGGCW		$1 \times 10^8$	5' - CTGGGCA - 3'
24	WWGCCGW		$2 \times 10^9$	5' - ATGCCGT - 3'
25	WWGGCCW		$9 \times 10^9$	5' - ATGGCCA - 3'
26	WWGCGCW		$3 \times 10^9$	5' - ATGCGCA - 3'
27	WWGCCCW		$1 \times 10^9$	5' - ATGCCCA - 3'

**Table I.1.** Library of imidazole-capped polyamides. Shown are the targeted sequence, binding association constant ( $K_a$ ), and sequence context for binding affinity determination for each polyamide (shown as ball-and-stick models).<sup>35</sup>





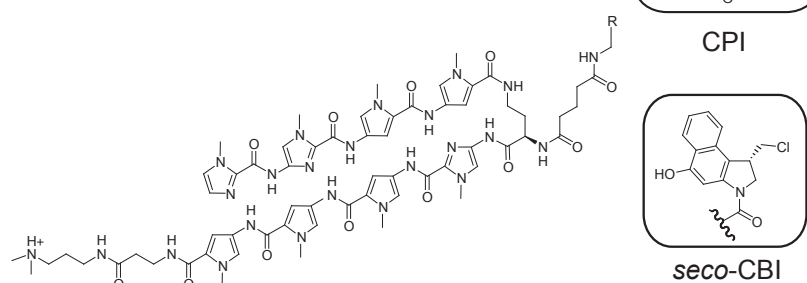
**Figure I.8.** Polyamides as regulators of gene expression in cell culture. **(A)** Schematic diagram of the VEGF promoter showing inhibition of HRE binding by HIF-1 (shown as HIF-1 $\alpha$ /HIF-1 $\beta$  heterodimer), binding sequence of the HRE enhancer shown with match polyamide **1**, and chemical structures and ball-and-stick models of match polyamide **1** and mismatch polyamide **2**. **(B)** Schematic diagram of the androgen receptor transcription complex, binding sequence of the consensus ARE targeted by match polyamide **3**, and chemical structures and ball-and-stick models of match polyamide **3** and mismatch polyamide **4**.

### Sequence-specific alkylation of DNA

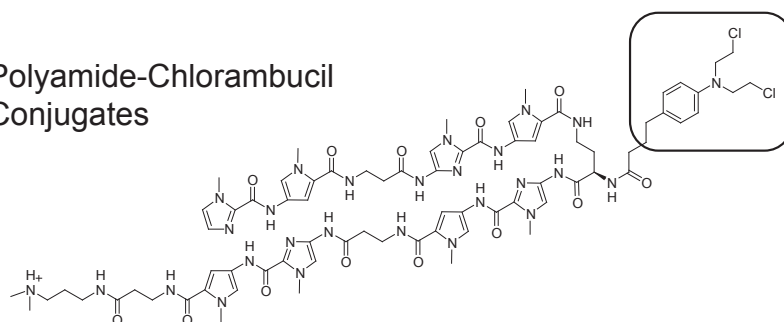
Polyamides have been used to regulate gene expression by binding to promoter and enhancer elements of DNA,<sup>40-42</sup> but do not prevent RNA polymerase elongation when bound to the coding region of genes.<sup>43</sup> DNA alkylators typically display limited sequence selectivity. Conjugation of an alkylating moiety to a sequence-specific molecule results in a molecule capable of covalent DNA attachment that arrests transcription.<sup>44</sup> An example of this is Tallimustine, which is a conjugate of distamycin A and the nitrogen mustard chlorambucil.<sup>45</sup>

Several types of alkylating agents have been tethered to polyamide hairpin turn units, including: 1-(chloromethyl)-5-hydroxy-1,2-dihydro-3*H*-benz[*e*]indole (*seco*-CBI),<sup>46</sup> cyclopropylpyrroloindole (CPI),<sup>47</sup> and chlorambucil (Chl) (Figure I.9).<sup>48</sup> Both *seco*-CBI and Chl conjugates showed alkylation at adenines proximal to the molecules' binding sites,<sup>46,48-50</sup> however polyamide-*seco*-CBI molecules were far more reactive at both match and mismatch sites than their chlorambucil counterparts. Hairpin polyamide-chlorambucil conjugates have been shown to target specific sequences of the HIV-1 promoter<sup>48</sup> and simian virus 40.<sup>51</sup>

### Polyamide-*seco*-CBI / CPI Conjugates



### Polyamide-Chlorambucil Conjugates



**Figure I.9.** Examples of various polyamide-alkylator conjugates. Shown are *seco*-CBI and CPI tethered to a polyamide targeting 5'-WWGGWCW-3' (top) and chlorambucil attached to a 2-β-2 polyamide targeting 5'-WWGCWGCW-4' (bottom).

### Polyamide synthesis

Synthesis methods for the polyamides described herein largely involve solid-phase techniques. Protocols have been developed using standard Boc coupling chemistry,<sup>52,53</sup> with cleavage via aminolysis followed by further derivatization on the turn and/or tail residues. Usage of either β-PAM or Kaiser oxime resin yields different C-terminal groups following cleavage with a primary amine, the former leaving a residual β-alanine, and the latter leaving none. While solid phase methods have proved useful for the rapid synthesis of a variety of polyamides on micromolar scales, they are inefficient for making larger quantities of a particular core. For this reason, solution-phase methodologies for gram-scale polyamide synthesis involving minimal chromatography have been recently developed.<sup>54</sup>

## Scope of this work

The work described in this thesis relates to the effects of the hairpin turn subunit on Py-Im polyamides, including binding affinity and specificity, and cellular uptake properties. It is divided into two sections: part II describes studies relating to  $\alpha$ -2,4-diaminobutyric acid as the hairpin subunit; part III describes effects of 3,4-diaminobutyric acid in the hairpin subunit. The appendix describes methods for genome-wide analysis of polyamide binding using expertise gained in part III.

Chapters IIA and IIB describe the identification and characterization of polyamides and polyamide-Chl conjugates linked by  $\alpha$ -2,4-diaminobutyric acid (also referred to as  $\alpha$ -diaminobutyric acid, or  $\alpha$ -DABA), with comparisons to analogous  $\gamma$ -2,4-diaminobutyric acid (also referred to as  $\gamma$ -diaminobutyric acid, or  $\gamma$ -DABA) containing molecules. Binding affinities and specificities for small libraries of  $\alpha$ - and  $\gamma$ -DABA linked polyamides were determined, and alkylation properties of the Chl conjugates explored. Effects of  $\alpha$ - and  $\gamma$ -DABA-linked molecules on colon carcinoma cell growth and on mice were evaluated. In chapter IIC,  $\alpha$ -DABA-linked polyamide-Chl conjugates targeting the coding region of histone H4c are studied *in vitro* and *in vivo* in an effort to affect chronic myelogenous leukemia.

Chapters IIIA and IIIB characterize polyamides containing various substituents in the  $\gamma$ -turn unit. In Chapter IIIA, polyamides containing (*R*) and (*S*)-3,4-diaminobutyric acid ( $\beta$ -amino- $\gamma$ -turn) in the hairpin turn are compared to analogues containing unsubstituted  $\gamma$ -DABA and  $\gamma$ -2,4-diaminobutyric acid (referred to as  $\alpha$ -amino- $\gamma$ -turn in part III) with respect to binding abilities and in cell culture. In Chapter IIIB, six-ring polyamides are employed to determine the binding affinities and specificities of  $\alpha$ -amino- $\gamma$ -turn, (*R*) and (*S*)- $\beta$ -amino- $\gamma$ -turn, and fluoro- and hydroxyl-turn polyamides at the turn position.

In the appendix, a polyamide-biotin conjugate containing a  $\beta$ -amino- $\gamma$ -turn subunit targeting the androgen response element is shown to affect androgen-responsive gene expression and used to capture chromatin following treatment in a prostate cancer cell

line to determine genome-wide binding of polyamides.

## References

1. Collins, F. S.; Lander, E. S.; Rogers, J.; Waterston, R. H. *Nature*. **2004**. 431, 931-945.
2. Pabo, C. O.; Sauer, R. T. *Annu. Rev. Biochem.* **1984**. 53, 293-321.
3. Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell*. **1992**. 71, 1223-1237.
4. Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature*. **1980**. 287, 755-758.
5. Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M.L. *Science*. **1982**. 216, 475-485.
6. Chen, F. E.; Huang, D. B.; Chen, Y. Q.; Ghosh, G. *Nature*. **1998**. 391, 410-413.
7. Pavletich, N. P.; Pabo, C. O. *Science*. **1991**. 252, 809-817.
8. Werner, M. H.; Clore, G. M.; Fisher, C. L.; Fisher, R. J.; Trinh, L.; Shiloach, J.; Gronenborn, A. M. *J. Biomol. NMR*. **1997**. 10, 317-328.
9. Juo, Z. S.; Chiu, T. K.; Leiberman, P. M.; Baikalov, I.; Berk, A. J.; Dickerson, R. E. *J. Mol. Biol.* **1996**. 261, 239-254.
10. Love, J. J.; Li, X. A.; Case, D. A.; Giese, K.; Grosschedl, R.; Wright, P. E. *Nature*. **1995**. 376, 791-795.
11. Kalben, A.; Pal, S.; Andreotti, A. H.; Walker, S.; Gange, D.; Biswas, K.; Kahne, D. *J. Am. Chem. Soc.* **2000**. 122, 8403-8412.
12. Vandyke, M. W.; Dervan, P. B. *Biochemistry*. **1983**. 22, 2373-2377.
13. Hou, M. H.; Robinson, H.; Gao, Y. G.; Wang, A. H. J. *Nucleic Acids Res.* **2004**. 32, 2214-2222.
14. Arcamone, F.; Nicoletti, V.; Penco, S.; Orezzi, P.; Pirelli, A. *Nature*. **1964**. 203, 1064-1065.
15. Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A. *Proc. Natl. Acad. Sci. U. S. A.* **1987**. 84, 8385-8389.

16. Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **1989.** 86, 5723-5727.
17. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M. *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **1999.** 55, 602-609.
18. Dervan, P. B. *Bioorg. Med. Chem.* **2001.** 9, 2215-2235.
19. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003.** 13, 284-299.
20. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C. *Nat. Struct. Biol.* **1998.** 5, 104-109.
21. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature.* **1998.** 391, 468-471.
22. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science.* **1998.** 282, 111-115.
23. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B. *Bioorg. Med. Chem.* **2003.** 11, 4333-4340.
24. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998.** 120, 6219-6226.
25. Wang, C. C. C.; Ellervik, U.; Dervan, P. B. *Bioorg. Med. Chem.* **2001.** 9, 653-657.
26. Chenoweth, D. M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2009.** 106, 13175-13179.
27. Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1992.** 89, 7586-7590.
28. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994.** 116, 7983-7988.
29. deClairac, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997.** 119, 7909-7916.
30. Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999.** 121, 1113-1120.
31. Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998.** 120, 1382-1391.
32. White, S.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997.** 119, 8756-8765.
33. Trauger, J. W.; Baird, E. E.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1996.** 118, 6160-6166.

34. Trauger, J. W.; Dervan, P. B. *Meth. Enzymol.* **2001.** 340, 450-466.
35. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B.; Puckett, J. W.; Wang, C. C.; Dervan, P. B. *Tetrahedron.* **2007.** 63, 6146-6151.
36. Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B. *Bioorg. Med. Chem.* **2002.** 10, 3313-3318.
37. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003.** 100, 12063-12068.
38. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004.** 32, 2802-2818.
39. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007.** 35, 363-370.
40. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2004.** 101, 16768-16773.
41. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *ACS Chem. Biol.* **2007.** 2, 561-571.
42. Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2007.** 104, 10418-10423.
43. Dervan, P. B.; Poulin-Kerstien, A. T.; Fechter, E. J.; Edelson, B. S. *Top. Curr. Chem.* **2005.** 253, 1-31.
44. Shinohara, K.; Sasaki, S.; Minoshima, M.; Bando, T.; Sugiyama, H. *Nucleic Acids Res.* **2006.** 34, 1189-1195.
45. Pezzoni, G.; Grandi, M.; Biasoli, G.; Capolongo, L.; Ballinari, D.; Giuliani, F. C.; Barbieri, B.; Pastori, A.; Pesenti, E.; Mongelli, N.; Spreafico, F. *Br. J. Cancer.* **1991.** 64, 1047-1050.
46. Chang, A. Y.; Dervan, P. B. *J. Am. Chem. Soc.* **2000.** 122, 4856-4864.
47. Bando, T.; Narita, A.; Saito, I.; Sugiyama, H. *J. Am. Chem. Soc.* **2003.** 125, 3471-3485.
48. Wurtz, N. R.; Dervan, P. B. *Chem. Biol.* **2000.** 7, 153-161.
49. Wyatt, M. D.; Lee, M.; Garbiras, B. J.; Souhami, R. L.; Hartley, J. A. *Biochemistry.* **1995.** 34, 13034-13041.

50. Wyatt, M. D.; Lee, M.; Hartley, J. A. *Anti-Cancer Drug Des.* **1997.** *12*, 49-60.
51. Wang, Y. D.; Dziegielewski, J.; Wurtz, N. R.; Dziegielewska, B.; Dervan, P. B.; Beerman, T. A. *Nucleic Acids Res.* **2003.** *31*, 1208-1215.
52. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996.** *118*, 6141-6146.
53. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B. *Bioorg. Med. Chem.* **2002.** *10*, 2767-2774.
54. Chenoweth, D. M.; Harki, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **2009.** *131*, 7175-7181.



## **II. Py-Im Polyamides Linked by 2,4-Diaminobutyric Acid in the Hairpin Turn Unit**

## Chapter IIA

### Unanticipated Differences Between $\alpha$ - and $\gamma$ -Diaminobutyric Acid-Linked Hairpin Polyamide-Alkylator Conjugates

*The text of this chapter was taken in part from a manuscript co-authored with Sherry M. Tsai, \* Peter B. Dervan, \* C. James Chou,<sup>‡</sup> and Joel M. Gottesfeld<sup>‡</sup> ( \* California Institute of Technology and <sup>‡</sup> The Scripps Research Institute)*

(Tsai, S.M.; Farkas, M.E.; Chou, C.J.; Gottesfeld, J.M.; Dervan, P.B. *Nucleic Acids Res.* **2007**, *35*, 307-316)

## Abstract

Hairpin polyamide-chlorambucil conjugates containing an  $\alpha$ -diaminobutyric acid ( $\alpha$ -DABA) turn moiety are compared to their constitutional isomers containing the well-characterized  $\gamma$ -DABA turn. Although the DNA-binding properties of unconjugated polyamides are similar, the  $\alpha$ -DABA conjugates display increased alkylation specificity and decreased rate of reaction. Treatment of a human colon carcinoma cell line with  $\alpha$ -DABA versus  $\gamma$ -DABA hairpin conjugates reveals only slight differences in toxicities while producing similar effects on cell morphology and G2/M stage cell cycle arrest. However, striking differences in animal toxicity between the two classes are observed. While mice treated with an  $\alpha$ -DABA hairpin polyamide do not differ significantly from control mice, the analogous  $\gamma$ -DABA hairpin is lethal. This dramatic difference from a subtle structural change would not have been predicted.

## Introduction

The ability to control gene expression through the use of DNA sequence-specific, cell-permeable molecules holds therapeutic promise. Based on the natural product distamycin A, pyrrole-imidazole polyamides are a class of molecules that can be programmed to bind a broad repertoire of DNA sequences with specificities comparable to naturally occurring DNA-binding proteins.<sup>1</sup> They have also been shown to permeate living cells, localize to the nuclei,<sup>2,3</sup> and regulate gene expression.<sup>4,5</sup>

Pairing rules have been established for minor groove recognition, whereby an *N*-methylpyrrole/*N*-methylimidazole (Py/Im) pair recognizes C•G, the reverse (Im/Py) recognizes G•C, and Py/Py specifies A•T or T•A.<sup>1,6</sup> The replacement of Py with  $\beta$ -alanine ( $\beta$ ) can enhance binding affinity by allowing the polyamide to reset its curvature with that of the DNA minor groove.<sup>7,8</sup> Polyamides may be linked via a turn moiety to give a hairpin structure which binds in the “forward” direction N→C with respect to the 5'→3' direction of the DNA strand.<sup>9</sup> In earlier studies, linkage of antiparallel three-ring subunits via a  $\gamma$ -aminobutyric acid moiety, resulting in a turn with three methylene units, was shown to enhance DNA binding affinity ~100-fold over unlinked subunits.<sup>10</sup> The  $\gamma$ -aminobutyric acid turn demonstrates selectivity for A•T or T•A base pairs over G•C and C•G.<sup>11</sup> Use of a glycine amino acid, with one methylene turn unit, resulted in a 3-fold reduction in affinity relative to the  $\gamma$ -aminobutyric acid. Subsequent NMR studies revealed that the glycine amino acid substituted polyamide bound DNA in an extended, rather than a hairpin, conformation.<sup>12,13</sup>

Substitution of the  $\gamma$ -aminobutyric acid with the chiral (*R*)-2,4-diaminobutyric acid ((*R*)- $\gamma$ -DABA) was later found to further increase DNA binding affinity ~10-fold.<sup>14</sup> In contrast, hairpin polyamides with the opposite enantiomer (*S*)- $\gamma$ -DABA result in an unfavorable steric clash with the DNA minor groove and decreased binding for the forward orientation. Remarkably, binding affinity to the opposite orientation, in which the polyamide binds C→N with respect to the 5'→3' direction of the DNA strand, is energetically favorable

due to relief of the steric clash.

Polyamides have been used to regulate gene expression by binding to promoter and enhancer elements, but they do not prevent RNA polymerase elongation when bound to the coding region of genes.<sup>1</sup> Conjugation of an alkylating moiety to a sequence-specific DNA-binding molecule would be expected to result in a molecule capable of covalent DNA attachment that could arrest transcription.<sup>15,16</sup> An early example of this is Tallimustine (FCE 24517), which is a conjugate of distamycin A and the nitrogen mustard chlorambucil.<sup>17</sup> When attached to a DNA minor groove binder, chlorambucil is capable of alkylation at the N3 position of an adenine proximal to the binding site.<sup>18,19</sup> Hairpin polyamide-chlorambucil conjugates have been demonstrated to target specific sequences of the HIV-1 promoter and simian virus 40.<sup>19,20</sup>

Recently, a hairpin polyamide-chlorambucil conjugate, synthesized at The Scripps Research Institute, was found to alkylate within the coding region of the histone H4c gene of the human colon carcinoma cell line SW620 and to block proliferation of these cells in culture.<sup>21</sup> In addition, this conjugate was found to block tumor growth in a SW620 xenograft mouse model without apparent toxicity. When we attempted to scale-up additional conjugate at Caltech, the newly-synthesized molecule was found to be highly toxic in athymic nude mice, contrary to prior experiments.<sup>22</sup> Subsequent studies showed that the two laboratories had utilized two different turn units in the synthesis of the conjugate: the conjugate synthesis at Scripps used Boc-D-Dab(Fmoc)-OH, while the conjugate synthesis at Caltech used Fmoc-D-Dab(Boc)-OH. The molecules are therefore constitutional isomers differing in the hairpin turn moiety. The original molecule made at Scripps continues to be called **1R-Chl**,<sup>22</sup> and the new molecule is designated **2R-Chl**. **1R-Chl** utilizes an  $\alpha$ -diaminobutyric acid ( $\alpha$ -DABA) turn unit, whereas **2R-Chl** employs the standard  $\gamma$ -DABA turn unit (Figure IIA.1). Linkage via the new  $\alpha$ -DABA turn unit results in one methylene group in the hairpin, whereas **2R-Chl** has three methylene groups. In addition, the linker between the polyamide core and the chlorambucil alkylator in **1R-Chl**

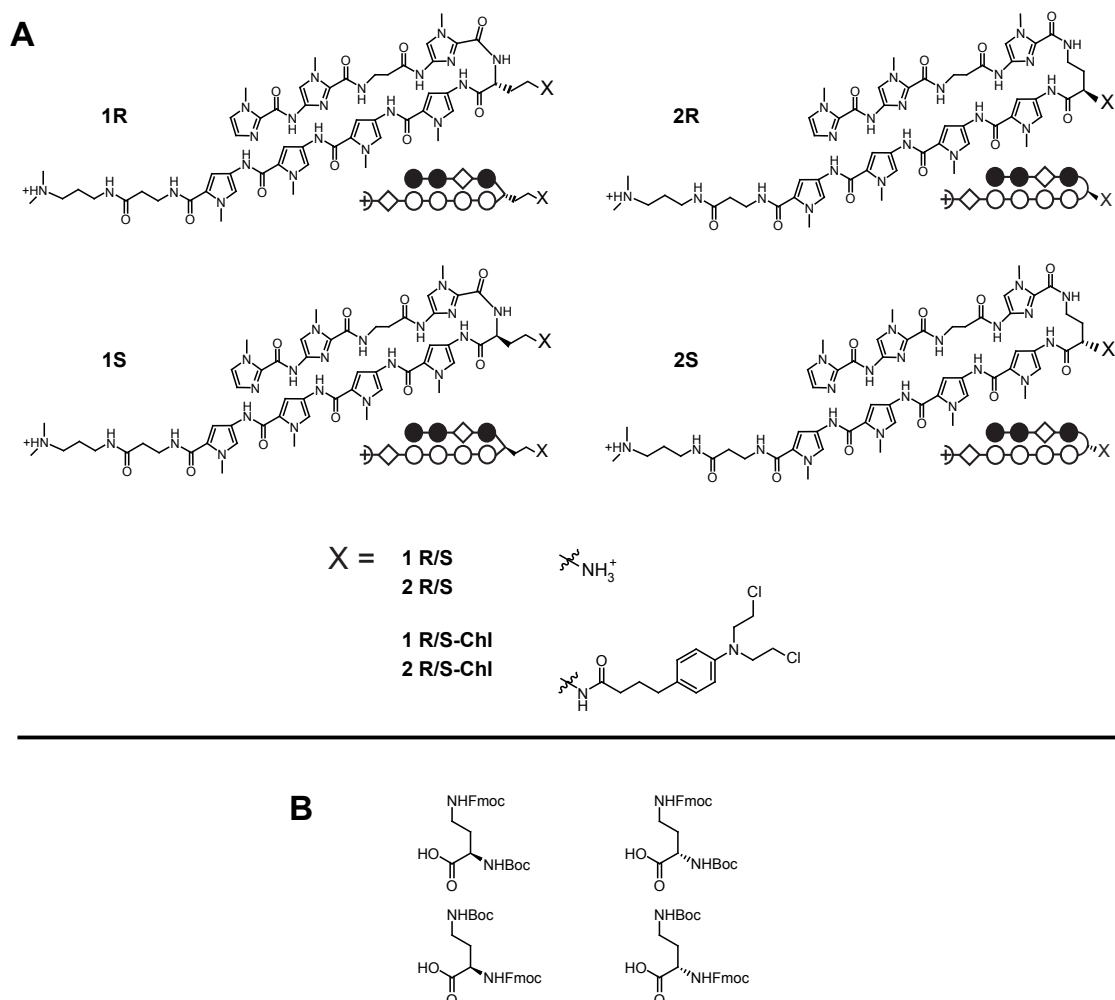
is lengthened by two methylene groups relative to **2R-Chl**.

Polyamides utilizing the  $\alpha$ -DABA turn moiety have not been studied previously, and the reasons for the unique biological properties displayed by **1R-Chl** are unknown. In the current study, we report the chemical and biological characterization of  $\alpha$ -DABA polyamides and compare them to the analogous  $\gamma$ -DABA polyamides. A series of four parent molecules and their corresponding chlorambucil conjugates was synthesized (Figure IIA.1A). Each molecule contains the heterocyclic sequence: ImIm $\beta$ Im-HT-PyPyPyPy- $\beta$ Dp (where HT = hairpin turn and Dp = 3-(dimethylamino)-propylamine). Molecules in series **1** contain the new  $\alpha$ -DABA turn, and molecules in series **2** utilize the  $\gamma$ -DABA turn. The *R* and *S* enantiomers of the parent polyamides and chlorambucil conjugates in each series have been generated for comparison. The DNA binding affinities of the parent molecules were measured by DNase I footprinting, and the alkylation profiles (sequence specificity and reactivity) of the polyamide-chlorambucil conjugates were established using thermal cleavage assays. The effects of polyamide treatment on SW620 cell morphology, proliferation, viability, and cell cycle profiles were compared using phase contrast microscopy, cell counting, trypan blue exclusion assay, and fluorescence-assisted cell sorting (FACS) analysis, respectively. Importantly, toxicity effects of the polyamide constitutional isomers on mice were determined.

## Results and Discussion

### Polyamide synthesis

Four hairpin polyamides were prepared by solid phase synthesis.<sup>23</sup> The hairpin polyamides differed in the turn substituents utilized: Boc-D-Dab(Fmoc)-OH (**1R**, **1R-Chl**), Boc-Dab(Fmoc)-OH (**1S**, **1S-Chl**), Fmoc-D-Dab(Boc)-OH (**2R**, **2R-Chl**), and Fmoc-Dab(Boc)-OH (**2S**, **2S-Chl**) (Figure IIA.1B). Conjugates were generated by coupling the HBTU-activated carboxylic acid of chlorambucil to the parent polyamide following



**Figure IIA.1.** Polyamide structures and syntheses. (A) Chemical and ball-and-stick structures for polyamides **1R**, **1S**, **2R**, **2S**, and their **Chl** conjugates. (B) Chemical structures of the four turn units used in the syntheses of the parent polyamides and polyamide-chlorambucil conjugates: Boc-D-Dab(Fmoc)-OH (top left, used in the syntheses of **1R** and **1R-Chl**); Boc-Dab(Fmoc)-OH (top right, used in the syntheses of **1S** and **1S-Chl**); Fmoc-D-Dab(Boc)-OH (bottom left, used in the syntheses of **2R** and **2R-Chl**); Fmoc-Dab(Boc)-OH (bottom right, used in the syntheses of **2S** and **2S-Chl**).

reverse-phase HPLC of the free amine and purified similarly.

### Plasmid design

The core polyamide structure was constructed to bind the sequence 5'-WWGGWGW-3', where W = A or T. In order to probe the affinity and sequence specificity of the parent polyamides, plasmid pMFST2 was designed with an insert containing a single match site

(‘M’ = 5’-TAGGTGT-3’) and two single base-pair mismatch sites (Figure IIA.2A). The mismatch sites consist of single G•C to C•G replacements at different locations within each binding site (‘A’ = 5’-TAGGTCT-3’, ‘B’ = 5’-TAGCTGT-3’). For the purpose of thermal cleavage assays, each binding site was designed with proximal adenines which the chlorambucil moiety would be expected to alkylate.<sup>19</sup> In addition to the designed sites, inherent in the plasmid structure is a single base-pair mismatch site (‘C’ = 5’-AGTGGTG-3’) for a polyamide binding in reverse (Figure IIA.3).

### DNA affinity and sequence specificity

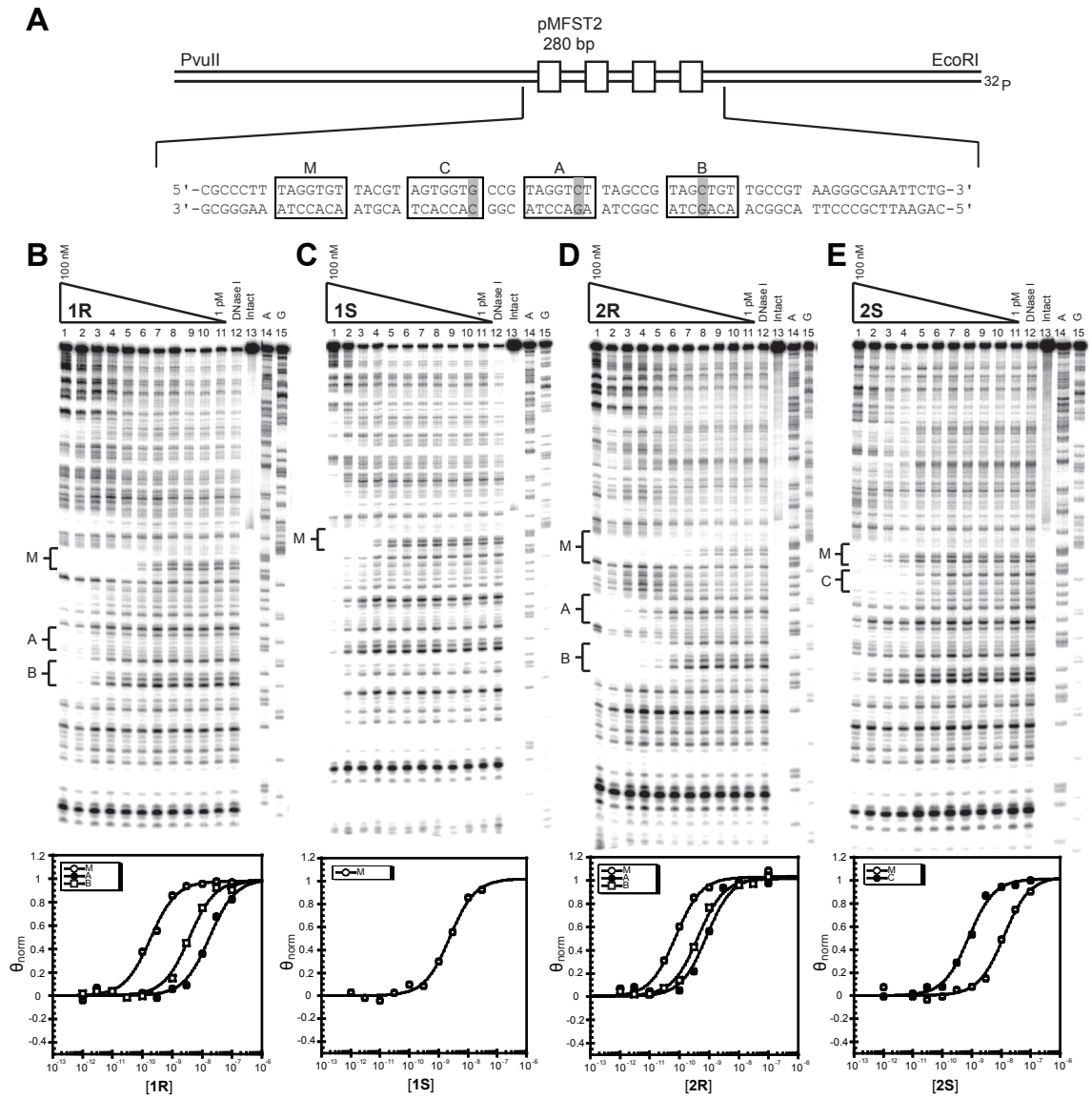
Quantitative DNase I footprint titrations to determine binding site affinities and specificities were conducted for parent polyamides **1R**, **1S**, **2R**, and **2S** on the 5’-labeled 280-base-pair PCR product of plasmid pMFST2 (Figure IIA.2 and Table IIA.1). Previous studies have indicated that attachment of the chlorambucil moiety does not significantly alter the DNA binding affinity of the parent polyamide.<sup>19</sup>

The  $\alpha$ -DABA polyamide **1R** bound the match site, M, with  $K_a = 6.7 \times 10^9 \text{ M}^{-1}$ , and it had a 94-fold specificity over single mismatch site A and 26-fold specificity over single mismatch site B (Figure IIA.2B). The enantiomer **1S** bound the match site M with  $K_a = 4.8 \times 10^8 \text{ M}^{-1}$ , a 14-fold decrease relative to **1R**. Binding affinities at the other sites were too low to be measured, and the polyamide was found to coat DNA at high concentrations (Figure IIA.2C).

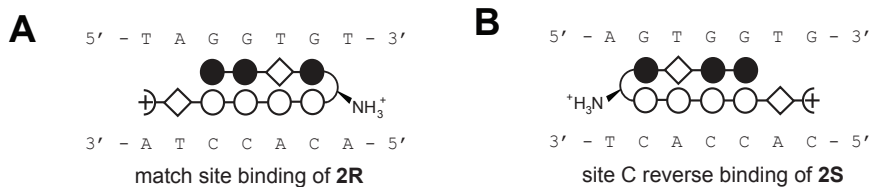
The  $\gamma$ -DABA polyamides **2R** and **2S** bound DNA in accordance with the expected orientation of the free amine.<sup>14</sup> **2R** was shown to bind the match and both single base-pair mismatch binding sites (Figure IIA.2D). The affinity of **2R** for site M was  $K_a = 1.6 \times 10^{10} \text{ M}^{-1}$ , with 15-fold and 6-fold preferences over mismatch sites A and B, respectively. The enantiomer, **2S**, bound the single base-pair mismatch reverse site C with  $K_a = 1.2 \times 10^9 \text{ M}^{-1}$ , a 21-fold enhancement over the forward match site M (Figure IIA.2E).

The binding isotherms for all four molecules were best-fit to a 1:1 binding model.





**Figure IIA.2.** Plasmid design and DNA binding properties of the parent polyamides. (A) Illustration of the designed portion of the EcoRI/PvuII restriction fragment derived from plasmid pMFST2. The four polyamide binding sites are indicated by boxes. Single base-pair mismatches are indicated by shaded regions. Site C is a mismatch site for a polyamide binding in reverse. (B-E) Quantitative DNase I footprint titration experiments for polyamides **1R** (B), **1S** (C), **2R** (D), **2S** (E) on the 280 bp, 5'-end-labeled PCR product of plasmid pMFST2: lanes 1-11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by its respective binding isotherms for the indicated sites (below).



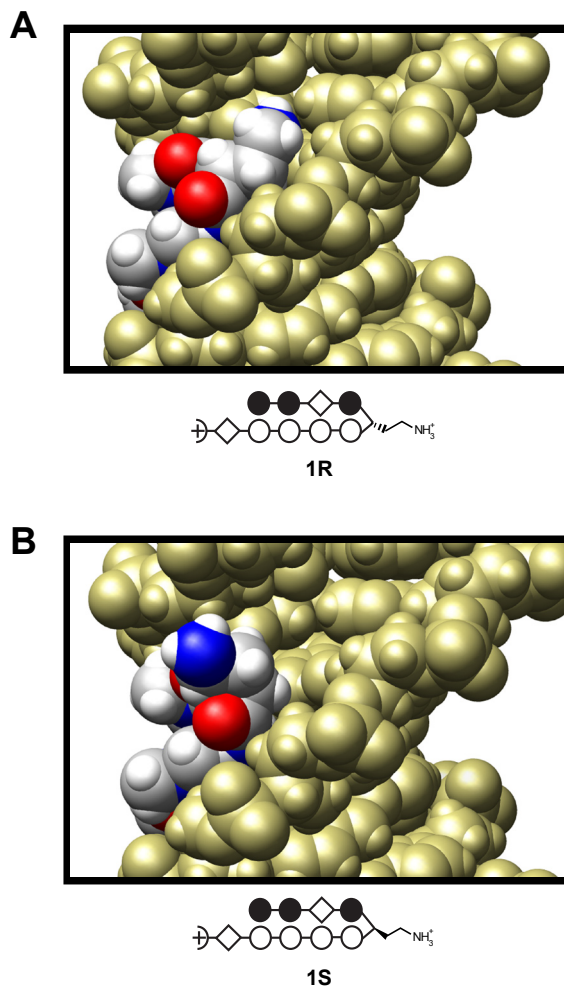
**Figure IIA.3.** Illustration of the parent polyamides (A) **2R** and (B) **2S** binding DNA.

Polyamide	5'-TAGGTGT-3'	5'-TAGGTCT-3'	5'-TAGCTGT-3'	5'-AGTGGTG-3'
<b>1R</b>	$6.7 (\pm 0.1) \times 10^9$	$7.1 (\pm 1.8) \times 10^7$	$2.6 (\pm 0.2) \times 10^8$	--
<b>1S</b>	$4.8 (\pm 0.9) \times 10^8$	--	--	--
<b>2R</b>	$1.6 (\pm 0.3) \times 10^{10}$	$1.1 (\pm 0.1) \times 10^9$	$2.9 (\pm 1.0) \times 10^9$	--
<b>2S</b>	$5.8 (\pm 2.1) \times 10^7$	--	--	$1.2 (\pm 0.2) \times 10^9$

**Table IIA.1.** Binding affinities ( $M^{-1}$ ) for parent polyamides. Equilibrium association constants reported are mean values from three DNase I footprint titration experiments; standard deviations are shown in parentheses.

Previous studies have found a glycine amino acid-linked polyamide is able to bind DNA in an extended dimeric conformation that can be observed via a larger binding site in a DNase I footprinting gel. We do not observe evidence of an extended conformation for either **1R** or **1S**, as the binding sites of both molecules are the same size as those of **2R** and **2S**. The substituted glycine favors a turn conformation.

The DNase I footprinting experiments reveal that the  $\alpha$ -DABA polyamide **1R** has a 2-fold reduced binding affinity relative to the  $\gamma$ -DABA polyamide **2R**. A contributing factor to the decreased binding observed for **1R** and **1S** could be a reduced ability of the molecules to fit into the minor groove. The shorter turn linkage between the two subunits in an  $\alpha$ -DABA polyamide may reduce its flexibility in conforming to the shape of the groove (Figure IIA.4). The energetics of  $\alpha$ -DABA hairpin polyamides appear less sensitive than the  $\gamma$ -DABA series to the *R* and *S* stereochemistry at the turn.



**Figure IIA.4.** Models of the binding of (A) **1R** and (B) **1S** to the match site of pMFST2. DNA atoms are indicated with gold; polyamide atoms are indicated as follows: carbon with gray, oxygen with red, nitrogen with blue, and hydrogen with white. Illustrations were rendered using Chimera.

### DNA alkylation specificity and time-dependence

To determine the alkylation properties of the polyamide-chlorambucil conjugates, thermal cleavage assays were performed on the 5'  $^{32}\text{P}$ -labeled 280-base-pair PCR product of pMFST2 (Figure IIA.5). *The alkylation profiles of 1R-Chl and 2R-Chl are dramatically different: 1R-Chl alkylates DNA more specifically than 2R-Chl and appears less reactive.* For the designed sequence, **1R-Chl** appears to specifically alkylate the adenine proximal to the turn at the match site M at concentrations up to 30 nM (Figure IIA.5A). It only

**Figure II A.5.** Alkylation specificities of the polyamide conjugates. Thermal cleavage assay experiments with (A) **1R-Chl**, (B) **2R-Chl**, (C) **1S-Chl**, and (D) **2S-Chl** on the 280 bp, 5'-end-labeled PCR product of pMFST2: lane 1, A reaction; lane 2, G reaction; lanes 3-13, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 14, intact DNA. Putative major sites of alkylation on the DNA fragment are indicated by arrows on the sequences adjacent to each gel.

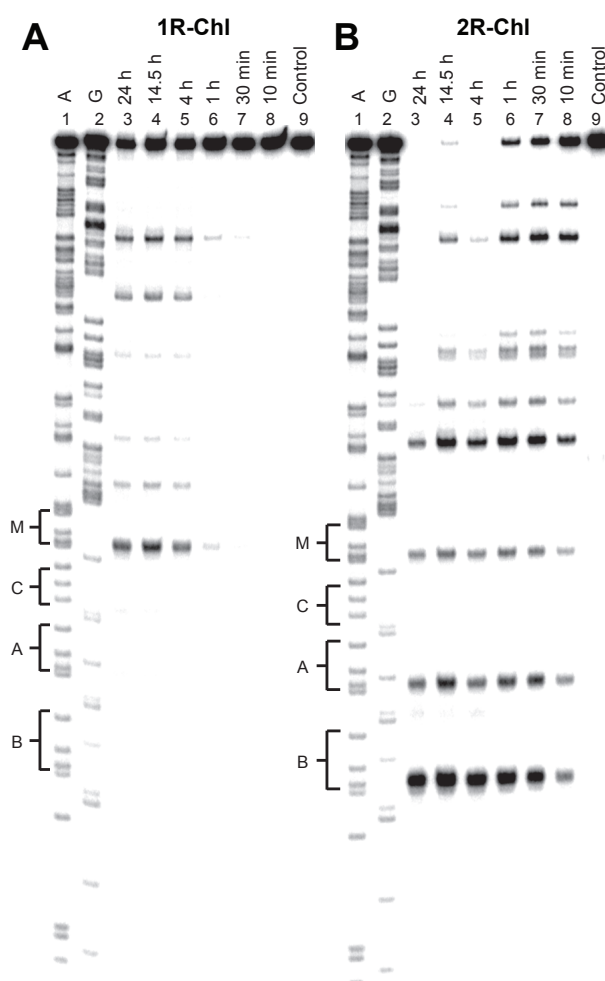
begins to alkylate at the designed mismatch binding sites at 100 nM concentration, and alkylation can be observed at concentrations as low as 1 nM. In contrast, **2R-Chl** is more promiscuous, alkylating at site M and both mismatch sites, A and B, at concentrations as low as 3 nM (Figure IIA.5B). Only at 1 nM concentration was **2R-Chl** observed to specifically alkylate at site M versus sites A and B.

The *S* enantiomers of the conjugates are similar to their *R* analogs in their alkylation specificities, but they are muted in activity. In contrast to **1R-Chl**, **1S-Chl** appears to specifically alkylate the guanine between sites M and C up to a concentration of 30 nM (Figure IIA.5C). Alkylation at N3 of guanine has been previously observed.<sup>18,21,24</sup> Based on the DNase I footprinting data for **1S**, the observed alkylation is probably due to binding of **1S-Chl** at the match site M. The **2S-Chl** alkylation profile at 100 nM is similar to that of **2R-Chl** at 10 nM, with the polyamide alkylating at the match site and both single base-pair mismatch sites (Figure IIA.5D). However, at 30 nM, **2S-Chl** is specific for the match site.

Outside of the designed pMFST2 insert, **1R-Chl** and **1S-Chl** alkylate fewer sites than **2R-Chl** and **2S-Chl**, respectively, further demonstrating their enhanced alkylation specificity. There are a number of forward-binding and reverse-binding single base-pair mismatch sites, as well as one reverse-binding match site, inherent in the labeled fragment of pMFST2 used in these studies. The major sites of alkylation observed for **1R-Chl** and **2R-Chl** outside of the designed plasmid insert can generally be attributed to forward-binding single base-pair mismatch sites, although the gel resolution is often too poor to definitively identify specific alkylation sites. The additional major alkylation sites observed for **2S-Chl** can be attributed to a combination of forward-binding and reverse-binding single base-pair mismatch sites, while the additional major alkylation site observed for **1S-Chl** may correspond to a reverse-binding single base-pair mismatch site.

We next studied the time-dependence of alkylation for **1R-Chl** and **2R-Chl** to quantitate their kinetic parameters (Figure IIA.6). At 500 nM polyamide and 37 °C, the

half-life of the labeled DNA (i.e. disappearance of full-length DNA) is 20 h for **1R-Chl** and 10 min for **2R-Chl**. Thus, **2R-Chl** alkylates DNA 120 times faster than **1R-Chl** under these conditions. Given the minimal structural differences between **1R-Chl** and **2R-Chl** and the similar binding affinities of their parent molecules, the dramatic difference in alkylation specificity and kinetics between hairpins with a common alkylator moiety would not have been predicted.



**Figure IIA.6.** Time-dependence of alkylation for polyamide conjugates. Thermal cleavage assay experiments with (A) 500 nM **1R-Chl** and (B) 500 nM **2R-Chl** on the 280 bp, 5'-end-labeled PCR product of pMFST2: lane 1, A reaction; lane 2, G reaction; lanes 3-8, equilibrations for 24 h, 14.5 h, 4 h, 1 h, 30 min, 10 min, respectively; lane 9, intact DNA.

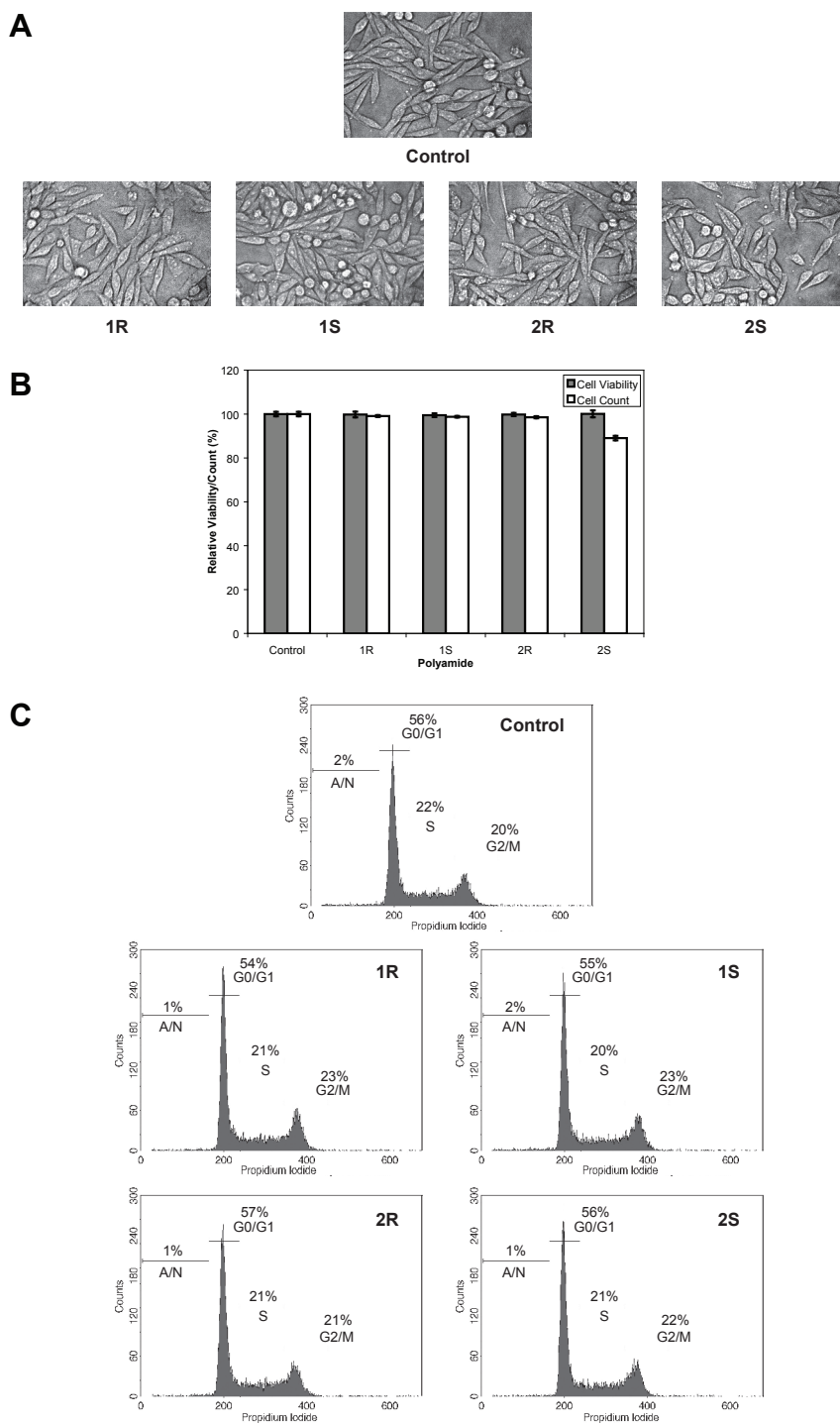
## Effects of polyamides on SW620 cells

**1R-Chl** treatment of cultured SW620 cells has been shown to induce a morphology change and arrest cell growth in the G2/M phase. In addition, the polyamide was found to be a cytostatic agent, decreasing cell proliferation without significantly affecting cell viability, as determined by trypan blue staining.<sup>21</sup> A two-hit mechanism for growth arrest by **1R-Chl** was suggested by experiments where a different polyamide-chlorambucil conjugate, which by itself had no effect on SW620 cell proliferation, was found to cause growth arrest after treatment of these cells with an siRNA directed against H4c mRNA.<sup>25</sup> Our model for the action of **1R-Chl** involves direct alkylation of the H4c gene, leading to a block in transcription and eventual depletion of H4 protein, opening up the genome for massive alkylation by **1R-Chl** and G2/M arrest.

We studied the effects of our polyamide library on SW620 cultured cells to determine whether the differences between the polyamides observed *in vitro* would translate to disparities in cell morphology, proliferation, viability, and cell-cycle profile. As expected, no effects were observed for the treatment of cells with the four parent polyamides (Figure IIA.7). Cells treated for 3 days with **1R-Chl** or **2R-Chl** show the same morphological change as previously established (Figure IIA.8A). Each conjugate also decreases cell proliferation to ~30% relative to control with only a small decrease in cell viability, as measured with trypan blue exclusion staining (Figure IIA.8B). FACS analysis using propidium iodide staining indicates both molecules arrest cells in the G2/M phase of growth (38% for **1R-Chl** and 34% for **2R-Chl**, compared to 20% for control; Figure IIA.8C).

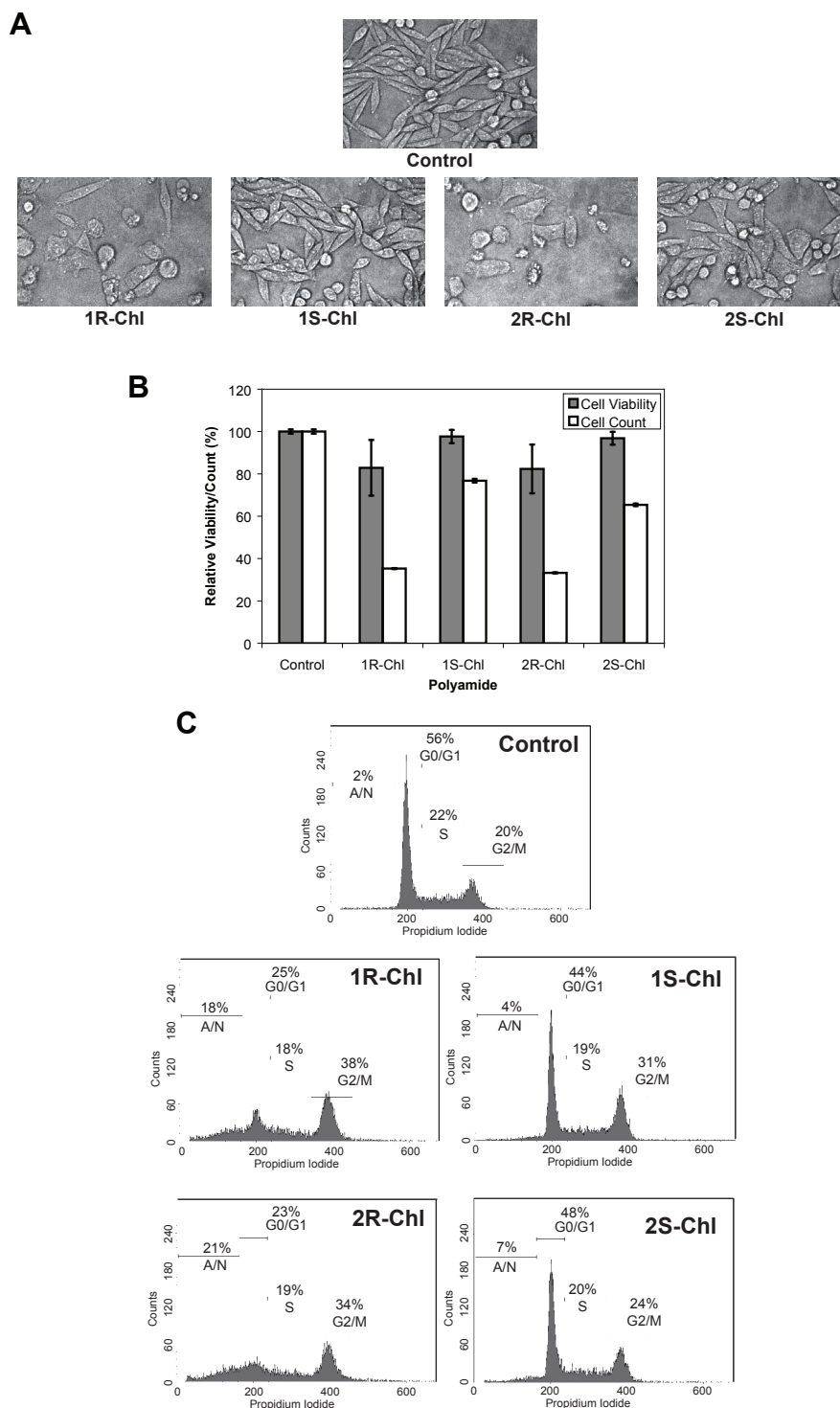
Interestingly, a less pronounced cellular effect is observed for cells treated with **1S-Chl** and **2S-Chl**. Although most cells appear identical to control, some of the cells treated with each of the *S* enantiomers display a morphology change (Figure IIA.8A). Cell viability is not significantly reduced, but proliferation decreases slightly (77% for **1S-Chl** and 65% for **2S-Chl** relative to control; Figure IIA.8B), indicating a slight cytostatic effect for these conjugates. This is also apparent in the FACS data for **1S-Chl**, where 31% of the





**Figure IIA.7.** Effects of parent polyamides on cultured SW620 cells. Cells were treated with 200 nM of the indicated polyamide for 3 days prior to analysis. **(A)** Representative phase microscopy images. **(B)** Normalized cell viability and proliferation values. Data shown are the averages of three samples; error bars indicate standard deviations. **(C)** Fluorescence-activated cell-sorting analysis of cells. Plots indicate cell numbers versus propidium iodide staining. Percentages of cells in G0/G1, S, G2/M, and apoptotic/necrotic cells (A/N) are indicated.





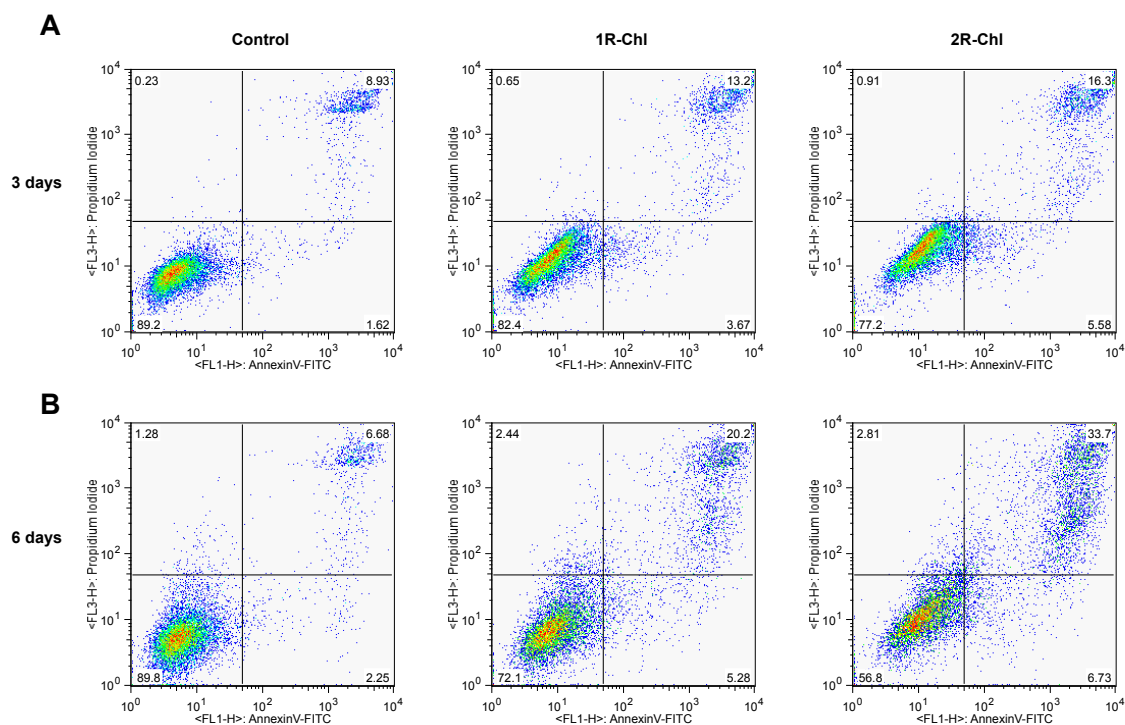
**Figure II A.8.** Effects of polyamide conjugates on cultured cells. Cultured SW620 cells were treated with 200 nM of the indicated polyamide for 3 days prior to analysis. (A) Representative phase microscopy images. (B) Normalized cell viability and proliferation values. Data shown are the averages of three samples; error bars indicate standard deviations. (C) Fluorescence-activated cell-sorting analysis of cells. Plots indicate cell numbers versus propidium iodide staining. Percentages of cells in G0/G1, S, G2/M, and apoptotic/necrotic cells (A/N) are indicated.

cells are found to be in the G2/M phase of growth. The greater population of apoptotic/necrotic cells following **2S-Chl** treatment may explain why significant growth arrest is not observed with these cells. We find that the trend of similar but muted effects of *S* relative to *R* enantiomers observed in the alkylation data appears to correlate to biological effects on SW620 cells. Previous studies have correlated the DNA interstrand crosslinking ability of alkylating agents with biological activity.<sup>26,27</sup> Since the *S* isomers are poor DNA alkylators and have minimal biological activity in SW620 cells, we chose not to investigate them further.

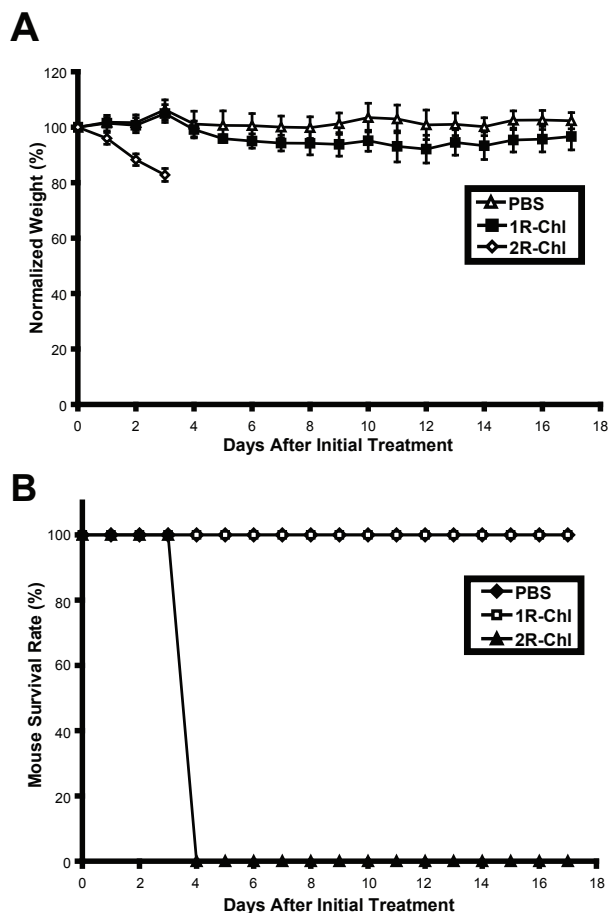
Additional experiments were performed with longer incubation times to determine if cells treated with **1R-Chl** or **2R-Chl** die by apoptosis (Figure IIA.9). SW620 cells were treated with 200 nM polyamide for either 3 or 6 days, with the 6-day experiment involving replacement of media with fresh media containing 200 nM polyamide on the third day. Comparison of the FACS analyses of these cells using annexin V-FITC and propidium iodide staining shows that after 6 days, **2R-Chl** treatment leads to greater cell death than **1R-Chl** treatment (33.7% versus 20.2%, respectively). However, a significant difference in apoptosis (i.e. cells that are annexin V positive and propidium iodide negative) between treatments with **1R-Chl** and **2R-Chl** is not observed (5.28% versus 6.73%, respectively).

### ***In vivo* effects of polyamides on BALB/c mice**

**1R-Chl** has been shown to arrest cancer growth in a SW620 xenograft nude mouse tumor model, whereas **2R-Chl** treatment was toxic to the mice.<sup>21</sup> The effects of **1R-Chl** and **2R-Chl** on a normal mouse strain were tested. Groups of five BALB/c mice were injected with PBS (control), 100 nmol **1R-Chl**, or 100 nmol **2R-Chl** every other day for up to 5 days. The mice were observed and weighed each day for up to 17 days (Figure IIA.10A). Mice treated with **1R-Chl** do not differ significantly from the control group. However, **2R-Chl** treatment of mice results in an average weight loss of 17% after 3 days and is found to be lethal to 100% of the mice after 4 days (Figure IIA.10B).



**Figure IIA.9.** Fluorescence-activated cell-sorting analysis of the effects of polyamide conjugates on SW620 cells. Cells were treated with 200 nM of the indicated polyamide or no polyamide (control) for (A) 3 days or (B) 6 days prior to analysis. For 6 day treatments, media was exchanged for fresh media containing 200 nM polyamide after 3 days. Cells were incubated with annexin V-FITC (AV) and propidium iodide (PI) prior to analysis. Percentages of cells that are AV-/PI- (bottom left quadrants; viable cells), AV+/PI- (bottom right quadrants; early apoptotic cells), AV-/PI+ (top left quadrants), and AV+/PI+ (top right quadrants; end-stage apoptotic/dead cells) are indicated.



**Figure IIA.10.** Effects of polyamide conjugates on BALB/c mice. Female BALB/c mice were treated with PBS (control), 100 nmols **1R-Chl**, or 100 nmols **2R-Chl** at 0, 2, and 4 (PBS and **1R-Chl** only) days. Five mice were treated for each experimental condition. **(A)** Average mouse weights, normalized to weight at day 0, over time. Error bars indicate standard deviations. **(B)** Survival rates for treated mice.

## Conclusion

We have shown that changing the turn unit of a hairpin pyrrole-imidazole polyamide-chlorambucil conjugate from the standard  $\gamma$ -diaminobutyric acid to an  $\alpha$ -diaminobutyric acid can result in dramatically different chemical and biological properties. Although unconjugated parent hairpin polyamides **1R** and **2R** bind DNA with comparable binding affinities, their corresponding chlorambucil conjugates have different alkylation profiles. The polyamide conjugate **1R-Chl**, which contains the (*R*)- $\alpha$ -diaminobutyric acid turn

unit, alkylates DNA much more specifically and at a decreased rate relative to **2R-Chl**, which contains the (*R*)- $\gamma$ -diaminobutyric acid turn. Strikingly, treatment of a common mouse strain with **1R-Chl** did not appear to have adverse effects on the mice, whereas treatment with **2R-Chl** was lethal. Although it is possible that the difference in toxicities between the two molecules is due to variance in absorption, distribution, metabolism, or excretion, given their structural similarities, it is not unreasonable to speculate the origin is the dramatic difference in alkylation specificity and reactivity. These results would not have been predicted *a priori* and indicate that hairpin polyamides containing the  $\alpha$ -diaminobutyric acid turn unit may be an important class of DNA-binding small molecules with interesting biological properties.

## Materials and Methods

### Polyamide Synthesis and Characterization

Polyamides were synthesized on solid phase on Boc- $\beta$ -Pam resin using Boc-protected monomers and dimers according to previously described protocols.<sup>23</sup> The following monomers/dimers were used in generating each of the molecules: Boc-Py-OBt (OBt = benzotriazol-1-yloxy), Boc- $\beta$ -Im-OH, and Im-Im-OH. Boc deprotection was conducted at room temperature for 30 min using 80% TFA/DCM prior to the first coupling reaction and after addition of each heterocycle. Carboxylic acids were activated with *N,N*-diisopropylethylamine (DIEA) and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) for 30 min at 37 °C. Monomer and dimer couplings were allowed to continue for 1 h (OBt esters) or 2 h (activated carboxylic acids). Polyamides were cleaved from resin with 3-(dimethylamino)-propylamine neat at 55 °C for ~18 h, and subsequently purified by reverse-phase high-performance liquid chromatography (HPLC).

Conjugation of polyamides to chlorambucil (4 eq) proceeded by activating the

carboxylic acid with HBTU (4 eq) and DIEA (excess) for 30 min at room temperature, after which the activation mixture was added to the polyamide. After allowing the reaction to continue for 1.5-2 h, products were purified by reverse-phase HPLC and immediately lyophilized.

The purity of all compounds were established by analytical HPLC and matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry.

**1R: ImImβIm-(R)<sup>H2N</sup>α-PyPyPyPy-Dp** Synthesized using Boc-D-Dab(Fmoc)-OH.

MALDI-TOF  $C_{54}H_{71}N_{22}O_{10}^+$  calculated  $[M+H]^+$ : 1188.28, found: 1188.02

**1S: ImImβIm-(S)<sup>H2N</sup>α-PyPyPyPy-Dp** Synthesized using Boc-Dab(Fmoc)-OH.

MALDI-TOF  $C_{54}H_{71}N_{22}O_{10}^+$  calculated  $[M+H]^+$ : 1188.28, found: 1187.99

**2R: ImImβIm-(R)<sup>H2N</sup>γ-PyPyPyPy-Dp** Synthesized using Fmoc-D-Dab(Boc)-OH.

MALDI-TOF  $C_{54}H_{71}N_{22}O_{10}^+$  calculated  $[M+H]^+$ : 1188.28, found: 1187.95

**2S: ImImβIm-(S)<sup>H2N</sup>γ-PyPyPyPy-Dp** Synthesized using Fmoc-Dab(Boc)-OH.

MALDI-TOF  $C_{54}H_{71}N_{22}O_{10}^+$  calculated  $[M+H]^+$ : 1188.28, found: 1187.77

**1R-Chl: ImImβIm-(R)<sup>Chl</sup>α-PyPyPyPy-Dp** Synthesized from **1R**. MALDI-TOF  $C_{68}H_{88}$

$Cl_2N_{23}O_{10}^+$  calculated  $[M+H]^+$ : 1474.48, found: 1474.98

**1S-Chl: ImImβIm-(S)<sup>Chl</sup>α-PyPyPyPy-Dp** Synthesized from **1S**. MALDI-TOF

$C_{68}H_{88}Cl_2N_{23}O_{10}^+$  calculated  $[M+H]^+$ : 1474.48, found: 1475.03

**2R-Chl: ImImβIm-(R)<sup>Chl</sup>γ-PyPyPyPy-Dp** Synthesized from **2R**. MALDI-TOF

$C_{68}H_{88}Cl_2N_{23}O_{10}^+$  calculated  $[M+H]^+$ : 1474.48, found: 1474.92

### **2S-Chl: ImImβIm-(S)<sup>Chl</sup>γ-PyPyPyPy-Dp** Synthesized from **2S**. MALDI-TOF

$C_{68}H_{88}Cl_2N_{23}O_{10}^+$  calculated  $[M+H]^+$ : 1474.48, found: 1474.73

### **Construction of Plasmid pMFST2**

Oligonucleotides were purchased from Integrated DNA Technologies. The plasmid pMFST2 was constructed by annealing the two oligonucleotides 5'-AGCTGCCGCCCTTTAGGTGTTACGTAGTGGTGCCGTAGGTCTTAGCCGTAGCTGTTGCCGTAAGG-GCGAATTCTGC-3' and 5'-GATCGCAGAATTCGCCCTTACGGCAACAGCTACGGCTAAGACCTACGGCACCACTACGTAACACCTAAAGGGCGC-3', followed by ligation into the BamHI/HindIII restriction fragment of pUC19 using T4 DNA ligase. The plasmid was then transformed into *E. Coli* JM109 competent cells. Ampicillin-resistant white colonies were selected from 25 mL Luria-Bertani agar plates containing 50 mg/mL ampicillin treated with XGAL and IPTG solutions, and grown overnight at 37 °C. Cells were harvested the following day, and purification of the plasmid was performed with a Wizard Plus Midiprep DNA purification kit (Promega). DNA sequencing of the plasmid insert was performed by the Sequence Analysis Facility at the California Institute of Technology.

### **Preparation of 5' <sup>32</sup>P-End-Labeled DNA**

The primer 5'-GAATTCGAGCTCGGTACCCGGG-3' was labeled at the 5'-end and subsequently used with the primer 3'-CAGCCCTTTGGACAGCACGGTC-5' to amplify plasmid pMFST2 as previously described.<sup>28</sup>

### **DNase I Footprint Titrations**

Polyamide equilibrations and DNase I footprint titrations were conducted on the 5'-end-labeled PCR product of pMFST2 according to standard protocols.<sup>28</sup> DNA was

incubated with polyamide conjugates or water (control) for 12 h at room temperature prior to reaction with DNase.

### **Thermal Cleavage Assays**

Thermal cleavage assay experiments were conducted on the 5'-end-labeled PCR product of pMFST2 as previously described.<sup>19</sup> DNA was incubated with polyamide conjugates or water (control) for 24 h at 37 °C prior to work-up, unless otherwise noted.

### **Cell Culture**

The cells used in this work were derived from the human colon adenocarcinoma cell line SW620 (purchased from ATCC) and were maintained in Leibovitz's L-15 medium (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (Cambrex), 10 mM HEPES buffer (Gibco), 1 mM sodium pyruvate (Gibco), and 1% antibiotic-antimycotic (Invitrogen). Cells were grown in a humidified 5% CO<sub>2</sub> atmosphere at 37 °C. For polyamide treatment experiments, cells were plated in 25 cm<sup>2</sup> cell culture flasks (Corning) at 600,000 cells/flask in 10 mL media. Flasks were incubated at 37 °C for 2.5 h prior to addition of polyamide stock solution or water (control). Three flasks were treated per condition. Cells were incubated for 3 days prior to analysis. Cell morphology was recorded by phase contrast microscopy with 40x magnification (Nikon Diaphot with attached Nikon D100 camera). Cells were then trypsinized for 7 min at 37 °C and recombined with their media. An aliquot was removed to determine cell viability and proliferation via trypan blue exclusion and cell count using a Vi-Cell XR cell viability analyzer (Beckman-Coulter). 100 images were obtained per flask (3 flasks/condition). The remaining cells from each treatment condition were then combined, pelleted, and washed with Dulbecco's PBS (Gibco), and fixed at -20 °C with 70% EtOH for 35 min. The fixed cells were then pelleted, washed with Dulbecco's PBS, re-pelleted, and suspended in a solution containing 0.1% Triton X-100, 0.2 mg/mL RNase A (Sigma), and 0.02 mg/mL propidium iodide (Sigma) in Dulbecco's PBS. Cells



were incubated overnight at 4 °C in the dark prior to FACS analysis (FACSCalibur; Becton Dickinson). Parameters for FL2 were adjusted to give histograms with G0/G1 peaks centered at 200 for each sample to compensate for differences in cell size.

### **Animal Experiments**

Female BALB/c mice were purchased from The Scripps Research Institute Division of Animal Resources. Mice 8-12 weeks of age were given injections of PBS (control) or 100 nmol polyamide/injection via the tail vein on days 0, 2, and 4, with the exception of **2R-Chl**. Doses of **2R-Chl** were given only on days 0 and 2 due to lethality by day 4. The mice were weighed and their behavior monitored and recorded daily. Experimental protocols were approved by The Scripps Institutional Animal Welfare Committee and were conducted in conformity with institutional guidelines, which are in compliance with national and international laws and policies.

### **Acknowledgements**

We thank David Alvarez for helpful discussions; David M. Chenoweth for molecular modeling; Kathy Klingensmith, Pastorcito Nieto, and Antonio Reyes from TSRI ICND Vivarium; and Dr. Kent Osborn DMV and Leslie Nielsen from TSRI Pathology lab for helpful discussions and assistance on animal care and protocols. We are grateful to the National Institutes of Health for research support (CA107311) and for a pre-doctoral NRSA training grant to M.E.F.

## References

1. Dervan, P. B.; Poulin-Kerstien, A. T.; Fechter, E. J.; Edelson, B. S. *Top. Curr. Chem.* **2005.** 253, 1-31.
2. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003.** 100, 12063-12068.
3. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004.** 32, 2802-2818.
4. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2004.** 101, 16768-16773.
5. Burnett, R.; Melander, C.; Puckett, J. W.; Son, L. S.; Wells, R. B.; Dervan, P. B.; Gottesfeld, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2006.** 103, 11497-11502.
6. Wemmer, D. E.; Dervan, P. B. *Curr. Opin. Struct. Biol.* **1997.** 7, 355-361.
7. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998.** 120, 6219-6226.
8. Wang, C. C. C.; Ellervik, U.; Dervan, P. B. *Bioorg. Med. Chem.* **2001.** 9, 653-657.
9. White, S.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997.** 119, 8756-8765.
10. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994.** 116, 7983-7988.
11. Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999.** 121, 1113-1120.
12. Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Nat. Struct. Biol.* **1996.** 3, 321-324.
13. de Clairac, R. P. L.; Seel, C. J.; Geierstanger, B. H.; Mrksich, M.; Baird, E. E.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1999.** 121, 2956-2964.
14. Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998.** 120, 1382-1391.
15. Shinohara, K.; Sasaki, S.; Minoshima, M.; Bando, T.; Sugiyama, H. *Nucleic Acids Res.* **2006.** 34, 1189-1195.
16. Bando, T.; Sugiyama, H. *Accounts Chem. Res.* **2006.** 39, 935-944.

17. Pezzoni, G.; Grandi, M.; Biasoli, G.; Capolongo, L.; Ballinari, D.; Giuliani, F. C.; Barbieri, B.; Pastori, A.; Pesenti, E.; Mongelli, N.; Spreafico, F. *Br. J. Cancer*. **1991**. *64*, 1047-1050.
18. Wyatt, M. D.; Lee, M.; Hartley, J. A. *Anti-Cancer Drug Des.* **1997**. *12*, 49-60.
19. Wurtz, N. R.; Dervan, P. B. *Chem. Biol.* **2000**. *7*, 153-161.
20. Wang, Y. D.; Dziegielewska, J.; Wurtz, N. R.; Dziegielewska, B.; Dervan, P. B.; Beerman, T. A. *Nucleic Acids Res.* **2003**. *31*, 1208-1215.
21. Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B. S.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M. *Chem. Biol.* **2004**. *11*, 1583-1594.
22. Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B. S.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M. *Chem. Biol.* **2006**. *13*, 339-339.
23. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**. *118*, 6141-6146.
24. Wyatt, M. D.; Lee, M.; Garbiras, B. J.; Souhami, R. L.; Hartley, J. A. *Biochemistry*. **1995**. *34*, 13034-13041.
25. Alvarez, D.; Chou, C. J.; Latella, L.; Zeitlin, S. G.; Ku, S.; Puri, P. L.; Dervan, P. B.; Gottesfeld, J. M. *Cell Cycle*. **2006**. *5*, 1537-1548.
26. Pawlak, K.; Pawlak, J. W.; Konopa, J. *Cancer Res.* **1984**. *44*, 4289-4296.
27. Sunters, A.; Springer, C. J.; Bagshawe, K. D.; Souhami, R. L.; Hartley, J. A. *Biochem. Pharmacol.* **1992**. *44*, 59-64.
28. Trauger, J. W.; Dervan, P. B. *Methods Enzymol.* **2001**. *340*, 450-466.

## Chapter IIB

### $\alpha$ -Diaminobutyric Acid-Linked Hairpin Polyamides

*The text of this chapter was taken in part from a manuscript co-authored with Sherry M. Tsai and Peter B. Dervan (California Institute of Technology)*

(Farkas, M.E.; Tsai, S.M; Dervan, P.B. *Bioorg. Med. Chem.* **2007**, *15*, 6927-6936)

## Abstract

A hairpin polyamide-chlorambucil conjugate linked by  $\alpha$ -diaminobutyric acid ( $\alpha$ -DABA) has been shown to have interesting biological properties in cellular and small animal models. Remarkably, this new class of hairpin polyamides has not been previously characterized with regard to energetics and sequence specificity. Herein we present a series of pyrrole-imidazole hairpin polyamides linked by  $\alpha$ -DABA and compare them to polyamides containing the standard  $\gamma$ -DABA turn unit. The  $\alpha$ -DABA hairpins have decreased binding affinities overall. However,  $\alpha$ -DABA polyamide-chlorambucil conjugates are sequence-specific DNA alkylators with increased specificities. Affinity cleavage studies of  $\alpha$ -DABA polyamide-EDTA conjugates confirmed their preference for binding DNA in a forward hairpin conformation. In contrast, an unsubstituted glycine-linked polyamide prefers to bind in an extended binding mode. Thus, substitution on the turn unit locks the  $\alpha$ -DABA polyamide into the forward hairpin binding motif.

## Introduction

Pyrrole-imidazole polyamides are a class of synthetic ligands that can be programmed to bind a broad repertoire of DNA sequences with affinities and specificities similar to DNA-binding proteins.<sup>1,2</sup> Eight-ring hairpin polyamides have been shown to localize to the nuclei of living cells,<sup>3,4</sup> and regulate transcription by interfering with transcription factor-DNA interfaces in the promoters of endogenous genes.<sup>5-9</sup> Conjugation of polyamides to the DNA alkylating agent chlorambucil produces molecules capable of sequence-specific covalent reaction with DNA,<sup>10</sup> previously demonstrated for targeted sequences in the HIV-1 promoter, simian virus 40, and the histone H4c gene.<sup>11-13</sup>

DNA recognition by polyamides depends on side-by-side, aromatic amino acid pairings in the minor groove. Pairing rules have been established for DNA base pair recognition: an *N*-methylpyrrole/*N*-methylimidazole (Py/Im) pair recognizes C•G, the reverse (Im/Py) specifies G•C, and a Py/Py pair targets both A•T and T•A.<sup>2</sup> For some sequences, replacement of Py with a flexible  $\beta$ -alanine ( $\beta$ ) residue can enhance polyamide binding affinity by relaxing ligand curvature, affording a more optimal fit in the minor groove of DNA.<sup>14</sup>

Although Py-Im polyamides will bind in the minor groove as anti-parallel dimers,<sup>15</sup> polyamides bind DNA optimally as a single oligomer in a hairpin structure.<sup>16</sup> Early studies demonstrated the optimal length for the turn element is three methylene units, via the use of  $\gamma$ -aminobutyric acid as the turn residue.<sup>16</sup> The polyamide hairpin binds in the ‘forward’ direction N→C with respect to the 5′→3′ direction of the DNA strand.<sup>17</sup> Utilization of the substituted turn (*R*)-2,4-diaminobutyric acid ( $\gamma$ -DABA) further enhances DNA binding affinity.<sup>18</sup> Notably, employment of a linkage containing one methylene unit – provided by the use of glycine as the turn amino acid – resulted in polyamides that bound DNA in an extended 2:1 binding mode, rather than as a 1:1 hairpin complex.<sup>16,19-21</sup> In contrast, the  $\gamma$ -aminobutyric acid linkage appears to prefer the U-shaped conformation rather than an extended dimeric conformation.<sup>20</sup>

We recently reported a new turn unit with remarkable properties in small animal models: (*R*)- $\alpha$ -diaminobutyric acid ( $\alpha$ -DABA), which is formally a substituted glycine linkage since it also includes a single methylene unit within the hairpin turn.<sup>22</sup> This initial study characterized an  $\alpha$ -DABA polyamide-chlorambucil conjugate and compared its biological activity to its  $\gamma$ -DABA analog. A dramatic difference in toxicity was observed for the two conjugates: normal mice treated with the  $\alpha$ -DABA polyamide-chlorambucil conjugate did not differ significantly from control, while the analogous  $\gamma$ -DABA conjugate was lethal.<sup>22</sup> Notably, only the  $\alpha$ -DABA conjugate has been demonstrated to arrest cancer growth in a SW620 xenograft nude mouse tumor model without apparent toxicity.<sup>13</sup>

Given the striking *in vivo* effects demonstrated by the  $\alpha$ -DABA polyamide conjugate, this new turn unit merits further study. In particular,  $\alpha$ -DABA polyamides appear to bind in a hairpin conformation and not as an extended dimer, as would be expected based on studies of glycine-linked polyamides.

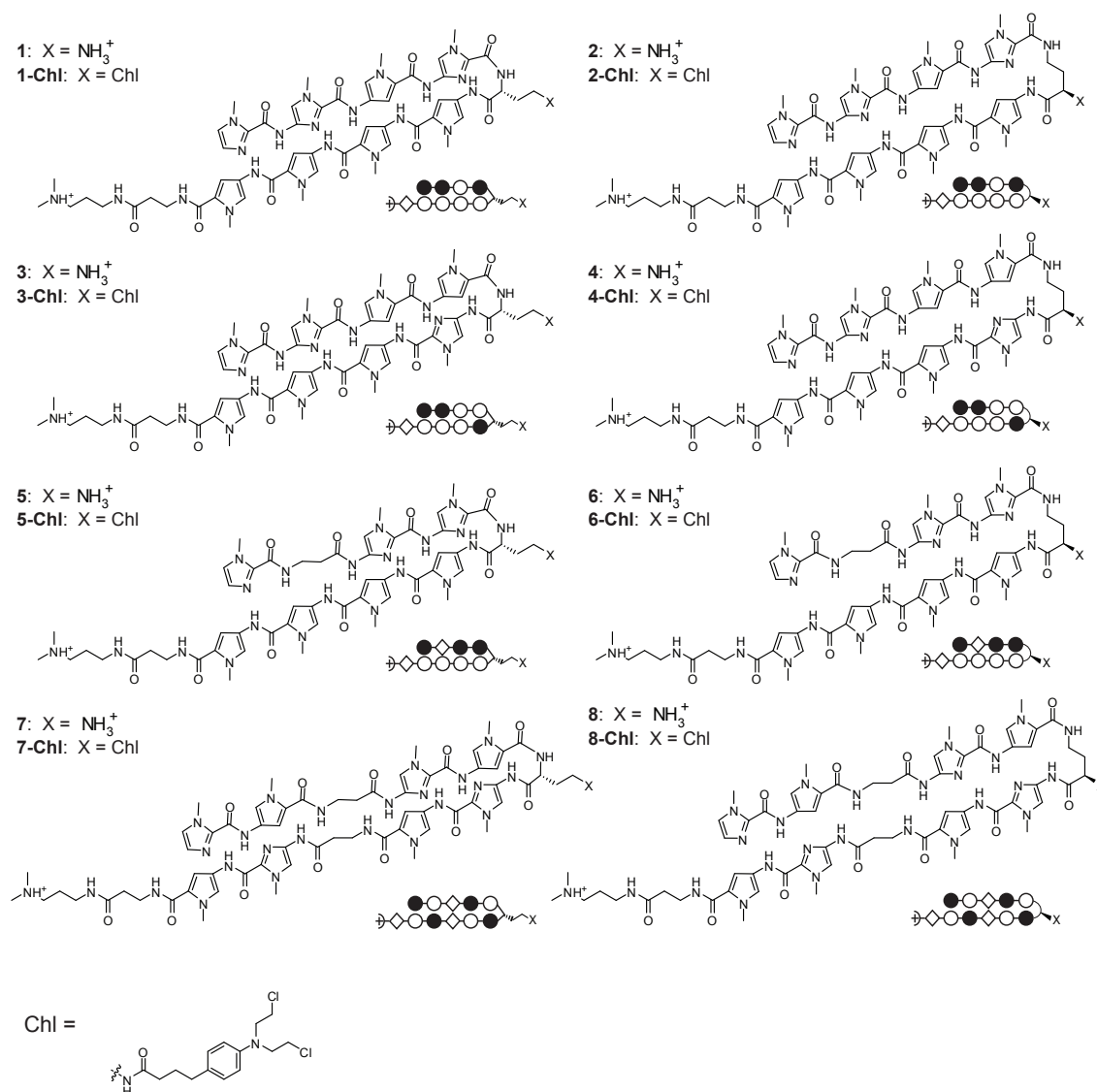
We report herein the characterization of  $\alpha$ -DABA polyamides. A series of  $\alpha$ -DABA polyamides and their chlorambucil conjugates were synthesized, along with their  $\gamma$ -DABA analogs. DNA binding properties of the parent molecules were established by DNase I footprinting titration experiments, and DNA alkylation properties of the chlorambucil conjugates were determined by thermal cleavage assay experiments. Importantly, we study the binding orientation of  $\alpha$ -DABA polyamides and address their ability to bind as hairpins by comparing them to their glycine-linked analogs.

## Results and Discussion

### Synthesis of the polyamide series

Four core ring sequences were selected for characterization (Figure IIB.1). Two standard eight-ring hairpin sequences, as well as a hairpin containing a  $\beta$ /Py pair and a 2- $\beta$ -2 configuration, were chosen. Eight parent hairpin polyamides were thus prepared: four

molecule pairs which code for different sequences according to the pairing rules. For each pair, one molecule was linked via the  $\alpha$ -DABA turn unit (polyamides **1**, **3**, **5**, and **7**), and the other was connected with the  $\gamma$ -DABA turn (polyamides **2**, **4**, **6**, and **8**). Polyamides were synthesized on Pam resin using standard solid phase methods.<sup>22,23</sup> The chlorambucil conjugates of the eight parent hairpin polyamides were also prepared based on previously published methods.<sup>22</sup>



**Figure IIB.1.** Chemical and ball-and-stick structures of polyamides **1-8** and their Chl conjugates. Shown are  $\alpha$ -diaminobutyric acid-linked hairpin polyamides (left) and  $\gamma$ -diaminobutyric acid-linked hairpin polyamides (right). The ball and stick symbols are defined as follows: an open circle denotes a pyrrole ring, a filled circle denotes an imidazole, and a diamond shape denotes  $\beta$ -alanine.



## Plasmid design

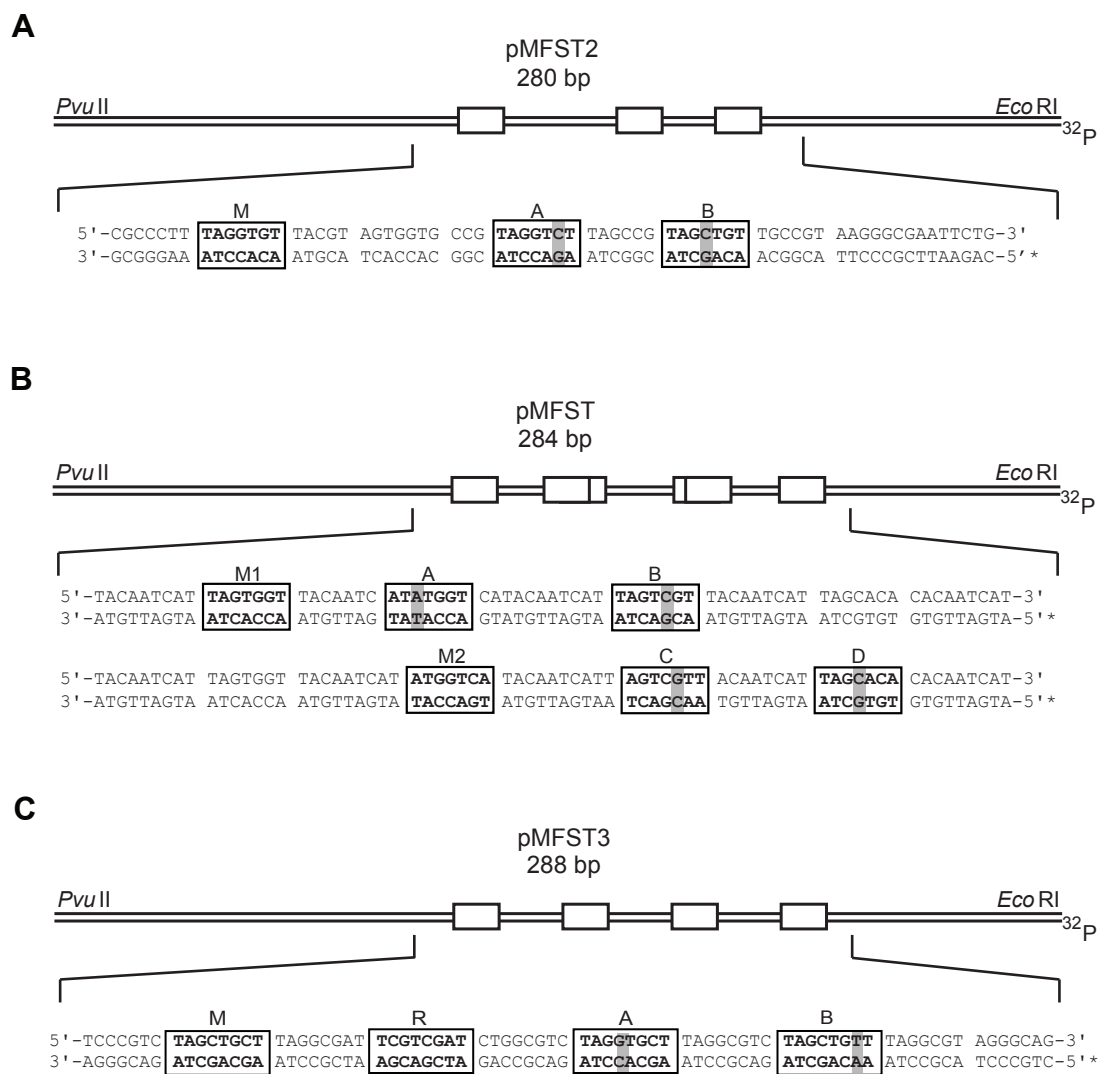
Polyamides **1** and **2** were designed to bind the sequence 5'-WWGGWGW-3' (where W = A or T). The plasmid pMFST2 was designed with an insert containing a single match site ('M' = 5'-TAGGTGT-3') and two single base pair mismatch sites ('A' = 5'-TAGGTCT-3' and 'B' = 5'-TAGCTGT-3') for polyamides **1** and **2** (Figure IIB.2A).<sup>22</sup>

A single plasmid, pMFST, was prepared to characterize polyamides **3-6** (Figure IIB.2B). Polyamides **3** and **4** were targeted to the sequence 5'-WWGGWCW-3', and polyamides **5** and **6** were designed to bind 5'-WWGWGGW-3'. The designed insert for pMFST contains match sites for each set of polyamides ('M1' = 5'-TAGTGGT-3', targeted to **5** and **6**, and 'M2' = 5'-ATGGTCA-3', targeted to **3** and **4**), as well as two single base pair mismatch sites for each ('A' = 5'-ATATGGT-3' and 'B' = 5'-TAGTCGT-3' for **5** and **6**, and 'C' = 5'-AACGACT-3' and 'D' = 5'-TAGCACA-3' for **3** and **4**).

The plasmid pMFST3 was used to characterize polyamides **7** and **8** (Figure IIB.2C) which target the sequence 5'-WWGCWGCW-3'. The insert for this plasmid contains a single match site ('M' = 5'-TAGCTGCT-3'), as well as a site for polyamide reverse binding ('R' = 5'-TCGTGAT-3'), and two single base pair mismatch sites ('A' = 5'-TAGGTGCT-3' and 'B' = 5'-TAGCTGTT-3'). The chlorambucil moiety is expected to alkylate at adenines proximal to the polyamide binding sites.<sup>11,22</sup> Thus, for the purposes of the thermal cleavage assays, adenines were placed next to the 3' end of the match sites and at least one of the mismatch sites for all plasmid inserts (Figure IIB.2).

## DNA binding affinity and sequence specificity

Quantitative DNase I footprinting titrations were performed with the parent polyamides to measure their binding site affinities and specificities (Figure IIB.3 and Table IIB.1). The  $\alpha$ -DABA polyamide **1** and its  $\gamma$ -DABA analog **2** were footprinted on the 5' <sup>32</sup>P-labeled 280 base pair PCR product of plasmid pMFST2 (Figure IIB.3A and 3B).  $\alpha$ -DABA polyamide **1** bound the match site M with  $K_a = 3.0 \times 10^8 \text{ M}^{-1}$  and the mismatch site A



**Figure IIB.2.** Illustration of the *EcoRI*/*PvuII* restriction fragments derived from plasmids (A) pMFST2, (B) pMFST, and (C) pMFST3. Polyamides 1 and 2 were characterized on pMFST2. Polyamides 3-6 were characterized on pMFST. Polyamides 7 and 8 were characterized on pMFST3. The designed polyamide binding sites are indicated by boxes. Single base pair mismatches are indicated by shaded regions. For pMFST, the designed insert is shown twice with binding sites for polyamides 3 and 4 (B, bottom) and polyamides 5 and 6 (B, top).

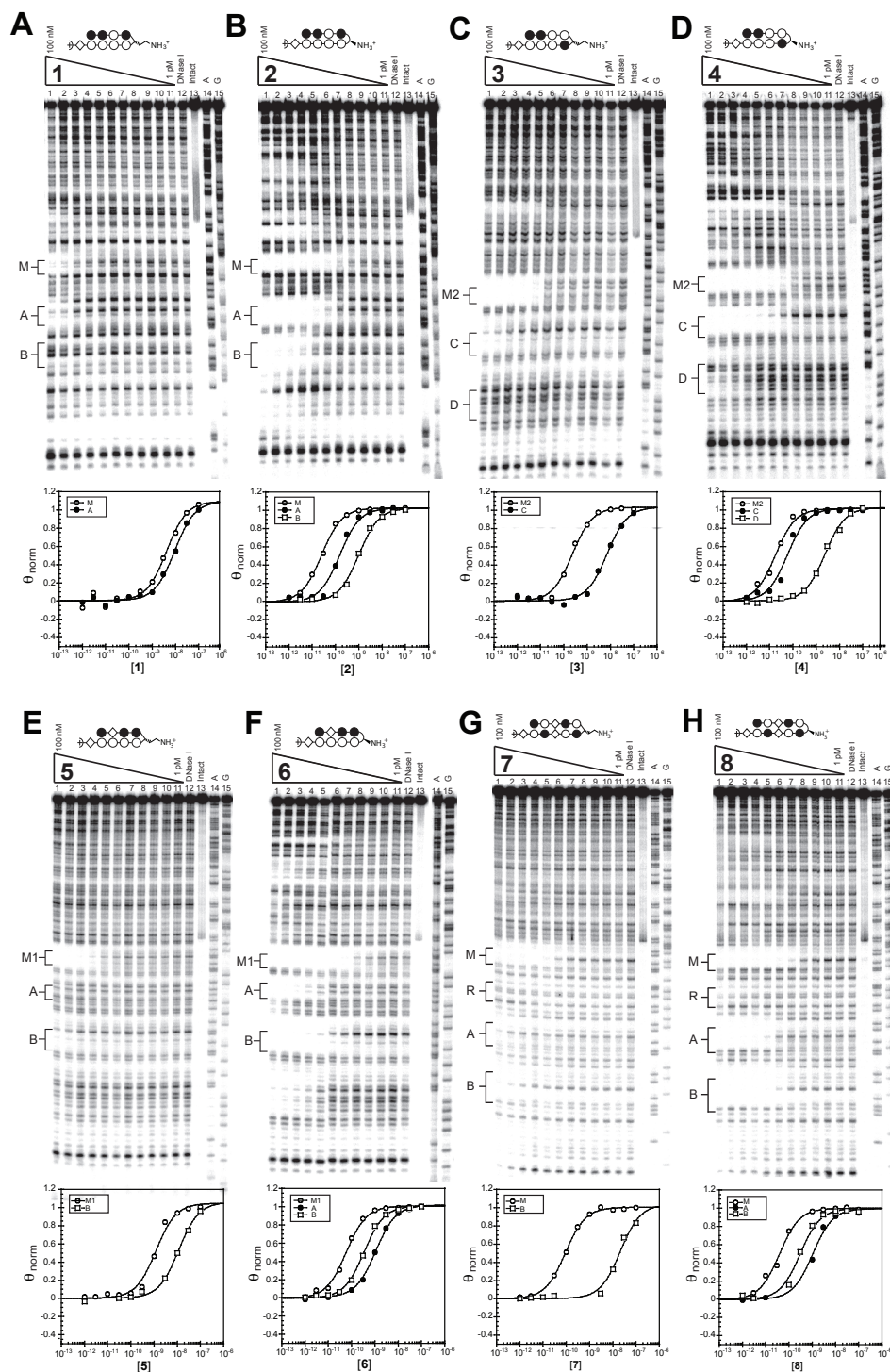
with  $K_a = 8.5 \times 10^7 \text{ M}^{-1}$  (Table IIB.1). Binding of **1** at the mismatch site B was not detected. In contrast, the  $\gamma$ -DABA polyamide **2** bound the match site M with  $K_a = 4.4 \times 10^{10} \text{ M}^{-1}$  and the mismatch sites A and B with  $K_a = 1.1 \times 10^{10} \text{ M}^{-1}$  and  $K_a = 1.1 \times 10^9 \text{ M}^{-1}$ , respectively. The  $\gamma$ -DABA polyamide **2** displays a 147-fold increased binding affinity over the  $\alpha$ -DABA polyamide **1** for their match site M.

The  $\alpha$ -DABA polyamides **3** and **5** and their  $\gamma$ -DABA analogs **4** and **6**, respectively, were footprinted on the 5'  $^{32}\text{P}$ -labeled 284 base pair PCR product of plasmid pMFST (Figure IIB.3C-3F).  $\alpha$ -DABA polyamide **3** bound its match site M2 with  $K_a = 4.4 \times 10^9 \text{ M}^{-1}$  and the mismatch site C with  $K_a = 1.5 \times 10^8 \text{ M}^{-1}$ . There was no detectable binding at mismatch site D. However, the analogous  $\gamma$ -DABA polyamide **4** bound the match site M2 with  $K_a = 4.9 \times 10^{10} \text{ M}^{-1}$  and the mismatch sites C and D with  $K_a = 1.8 \times 10^{10} \text{ M}^{-1}$  and  $K_a = 3.9 \times 10^8 \text{ M}^{-1}$ , respectively. The  $\alpha$ -DABA polyamide **3** binds its match site with 11-fold decreased affinity relative to the analogous  $\gamma$ -DABA polyamide **4**.

$\alpha$ -DABA polyamide **5** bound its match site M1 with  $K_a = 8.0 \times 10^8 \text{ M}^{-1}$  and its mismatch site B with  $K_a = 6.2 \times 10^7 \text{ M}^{-1}$ . Binding was not observed at mismatch site A. In comparison, the  $\gamma$ -DABA polyamide **6** binds the match site M1 with  $K_a = 4.3 \times 10^{10} \text{ M}^{-1}$  and the mismatch sites A and B with  $K_a = 9.9 \times 10^8 \text{ M}^{-1}$  and  $K_a = 6.2 \times 10^9 \text{ M}^{-1}$ , respectively. The  $\alpha$ -DABA polyamide **5** shows a 54-fold decreased affinity for its match site compared to the analogous  $\gamma$ -DABA polyamide **6**.

$\alpha$ -DABA polyamide **7** bound the match site M with  $K_a = 1.3 \times 10^{10} \text{ M}^{-1}$  and its mismatch site B with  $K_a = 1.0 \times 10^8 \text{ M}^{-1}$ . There was no detectable binding at mismatch site A. The analogous  $\gamma$ -DABA polyamide **8** binds the match site M with  $K_a = 2.9 \times 10^{10} \text{ M}^{-1}$  and mismatch sites A and B with  $K_a = 2.2 \times 10^9 \text{ M}^{-1}$  and  $K_a = 5.2 \times 10^9 \text{ M}^{-1}$ , respectively. The  $\alpha$ -DABA polyamide **7** shows a 2.2-fold decreased affinity for its match site compared to the  $\gamma$ -DABA polyamide **8**.

These results indicate that  $\alpha$ -DABA polyamides are proficient DNA binders, with decreased affinity and comparable or improved specificity relative to  $\gamma$ -DABA molecules.



**Figure IIB.3.** Quantitative DNase I footprinting titration experiments for polyamides (A) **1** and (B) **2** on the 280 base pair, 5' end-labeled PCR product of plasmid pMFST2; for polyamides (C) **3**, (D) **4**, (E) **5**, and (F) **6** on the 284 base pair, 5' end-labeled PCR product of plasmid pMFST; and for polyamides (G) **7** and (H) **8** on the 288 base pair, 5' end-labeled PCR product of plasmid pMFST3: lanes 1-11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by its respective binding isotherms (below).

The decrease in affinity is dependent on the polyamide core ring sequence. Varying the heterocyclic order of polyamides has also been shown to affect binding affinities in  $\gamma$ -DABA molecules.<sup>24</sup>

Polyamide	Match Site	$K_a$ ( $M^{-1}$ )	Mismatch Site	$K_a$ ( $M^{-1}$ )	Mismatch Site	$K_a$ ( $M^{-1}$ )
<b>1</b>	5'-TAGGTGT-3'	$3.0 (\pm 1.3) \times 10^8$ [147]	5'-TAGGT <b>T</b> C-3'	$8.5 (\pm 3.1) \times 10^7$ [129]	5'-TAG <b>C</b> TGT-3'	--
<b>2</b>	5'-TAGGTGT-3'	$4.4 (\pm 0.4) \times 10^{10}$	5'-TAGGT <b>T</b> C-3'	$1.1 (\pm 0.8) \times 10^{10}$	5'-TAG <b>C</b> TGT-3'	$1.1 (\pm 0.3) \times 10^9$
<b>3</b>	5'-ATGGTCA-3'	$4.4 (\pm 0.6) \times 10^9$ [11]	5'-AAC <b>G</b> ACT-3'	$1.5 (\pm 0.3) \times 10^8$ [120]	5'-TAG <b>C</b> ACA-3'	--
<b>4</b>	5'-ATGGTCA-3'	$4.9 (\pm 0.7) \times 10^{10}$	5'-AAC <b>G</b> ACT-3'	$1.8 (\pm 0.3) \times 10^{10}$	5'-TAG <b>C</b> ACA-3'	$3.9 (\pm 1.0) \times 10^8$
<b>5</b>	5'-TAGTGGT-3'	$8.0 (\pm 0.6) \times 10^8$ [54]	5'-ATATGGT-3'	--	5'-TAGT <b>C</b> GT-3'	$6.2 (\pm 3.0) \times 10^7$ [100]
<b>6</b>	5'-TAGTGGT-3'	$4.3 (\pm 0.3) \times 10^{10}$	5'-ATATGGT-3'	$9.9 (\pm 0.3) \times 10^8$	5'-TAGT <b>C</b> GT-3'	$6.2 (\pm 0.2) \times 10^9$
<b>7</b>	5'-TAGCTGCT-3'	$1.3 (\pm 0.2) \times 10^{10}$ [2.2]	5'-TAGGTGCT-3'	--	5'-TAGCTGTT-3'	$1.0 (\pm 0.7) \times 10^8$ [52]
<b>8</b>	5'-TAGCTGCT-3'	$2.9 (\pm 0.3) \times 10^{10}$	5'-TAGGTGCT-3'	$2.2 (\pm 0.3) \times 10^9$	5'-TAGCTGTT-3'	$5.2 (\pm 0.5) \times 10^9$

**Table IIB.1.** Binding affinities ( $M^{-1}$ ) for parent polyamides on plasmid pMFST2 (**1-2**), plasmid pMFST (**3-6**), and pMFST3 (**7-8**). Equilibrium association constants reported are mean values from three DNase I footprinting titration experiments. Standard deviations are shown in parentheses. Bolded numbers in brackets indicate fold decrease in match site binding affinity for  $\alpha$ -DABA-linked versus  $\gamma$ -DABA-linked polyamides. Mismatch base pairs are indicated in bold.

We previously reported the  $\alpha$ -DABA polyamide of sequence ImIm $\beta$ Im-( $R$ )<sup>H2N</sup> $\alpha$ -PyPyPyPy- $\beta$ Dp (where Dp = 3-(dimethylamino)-propylamine) binds the match site M of pMFST2 with  $K_a = 6.7 \times 10^9 M^{-1}$ .<sup>22</sup> Polyamide **1** has an identical structure except for the replacement of a  $\beta$  residue in the core sequence with a Py, and it binds the same match site M of pMFST2 with 22-fold decreased affinity. Polyamide **5** also has the same structure except for the movement of the  $\beta$  residue from the third position in the sequence to the second, and it binds its designed match site with 8-fold decreased affinity. These results suggest the presence and position of the  $\beta$  residue is important in determining the DNA binding affinities of  $\alpha$ -DABA polyamides. We hypothesize that the shortened  $\alpha$ -DABA turn unit requires the flexibility provided by the  $\beta$  to optimize binding.

## DNA alkylation properties

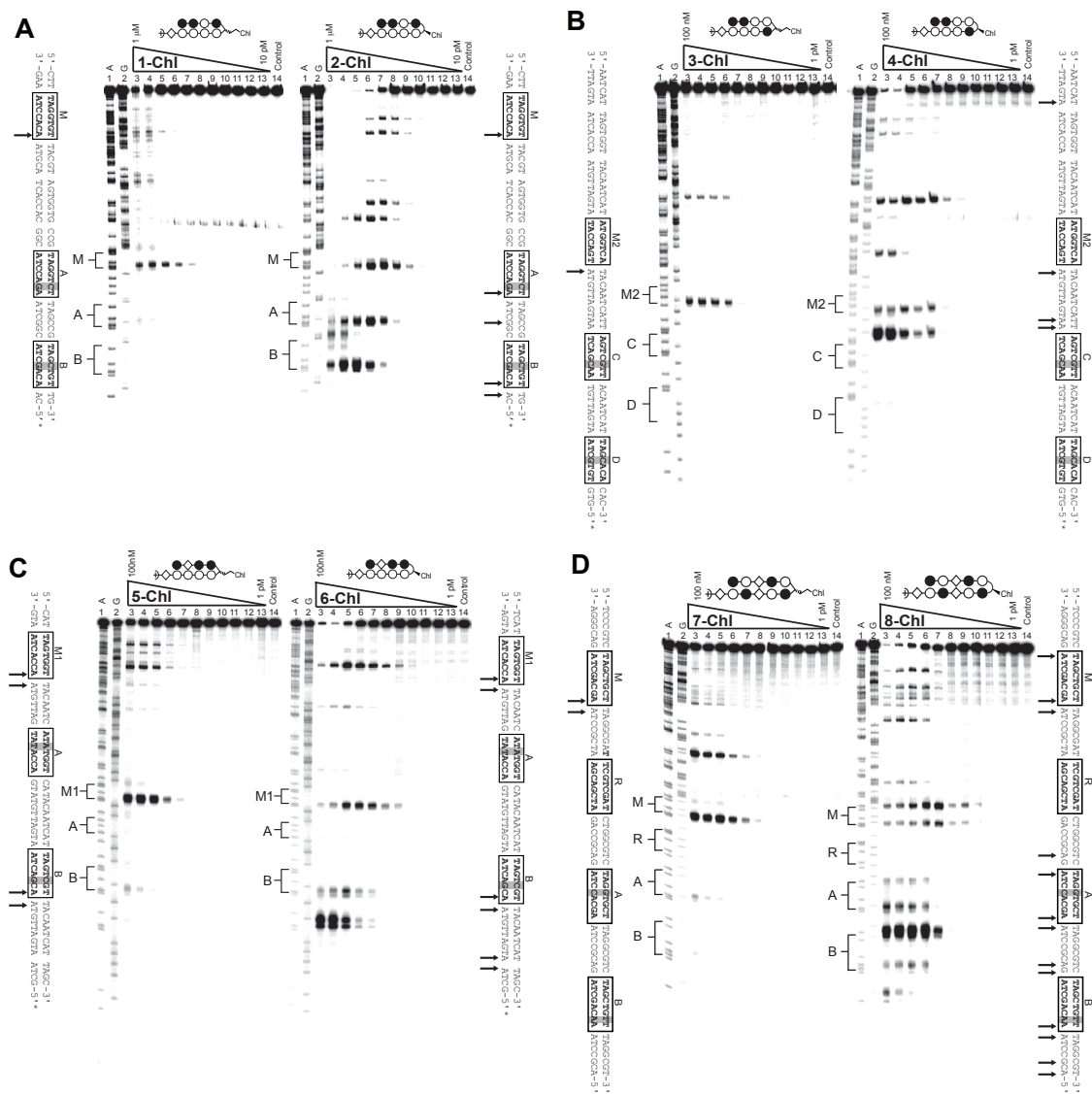
Thermal cleavage assays were performed with the polyamide-chlorambucil conjugates to determine their DNA alkylation properties (Figure IIB.4). Previous studies have shown conjugates to alkylate at N3 of adenines proximal to DNA binding sites.<sup>11,13</sup> Attachment of the chlorambucil moiety to the parent polyamide does not appear to affect DNA binding properties.<sup>11</sup>

The  $\alpha$ -DABA conjugate **1-Chl** and its  $\gamma$ -DABA analogue **2-Chl** were assayed on the 5' <sup>32</sup>P-labeled 280 base pair PCR product of plasmid pMFST2 (Figure IIB.4A). Alkylation was only observed in the designed insert of pMFST2 at the match site M for the  $\alpha$ -DABA conjugate **1-Chl**, whereas the  $\gamma$ -DABA conjugate **2-Chl** is more promiscuous, alkylating at the match site M and both mismatch sites A and B, as well as at additional unexpected sites.  $\gamma$ -DABA conjugate **2-Chl** also alkylates at lower concentrations than conjugate **1-Chl** (Figure IIB.4A, lanes 9).

Conjugates **3-Chl** and **4-Chl** were assayed on the 5' <sup>32</sup>P-labeled 284 base pair PCR product of plasmid pMFST (Figure IIB.4B). While parent polyamide **3** binds at site M2 and site C,  $\alpha$ -DABA conjugate **3-Chl** only alkylates at match site M2 in the designed insert. Alkylation is observed at both the match site M2 and the mismatch site C for the  $\gamma$ -DABA conjugate **4-Chl**. Although the parent polyamide **4** binds at mismatch site D, alkylation would not be expected to occur there because an A•T base pair is not located adjacent to the turn unit upon conjugate binding. Again, alkylation is observed at lower concentrations for the  $\gamma$ -DABA conjugate **4-Chl** than the  $\alpha$ -DABA conjugate **3-Chl** (Figure IIB.4B, lanes 7).

Conjugates **5-Chl** and **6-Chl** were also assayed on the PCR product of plasmid pMFST (Figure IIB.4C). For the  $\alpha$ -DABA conjugate **5-Chl**, strong alkylation is only observed at match site M1 in the designed insert, with minor alkylation occurring at mismatch site B. In contrast, the  $\gamma$ -DABA conjugate **6-Chl** strongly alkylates at both match site M1 and mismatch site B, as well as at other unexpected sites. As previously,





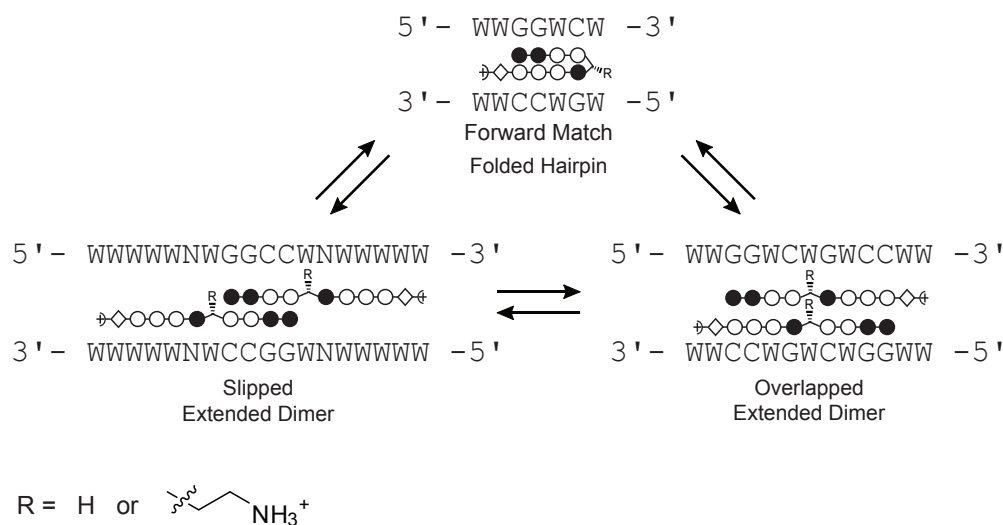
**Figure IIB.4.** (A) Thermal cleavage assay experiments with polyamide-chlorambucil conjugates (left) **1-Chl** and (right) **2-Chl** on the 280 base pair, 5' end-labeled PCR product of plasmid pMFST2: lane 1, A reaction; lane 2, G reaction; lanes 3-13, 1  $\mu$ M, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, and 10 pM polyamide, respectively; lane 14, intact DNA. (B-D) Thermal cleavage assay experiments with polyamide-chlorambucil conjugates **3-Chl** (B, left) and **4-Chl** (B, right), and **5-Chl** (C, left) and **6-Chl** (C, right) on the 284 base pair, 5' end-labeled PCR product of plasmid pMFST2, and **7-Chl** (D, left) and **8-Chl** (D, right) on the 288 base pair, 5' end-labeled plasmid pMFST3: lane 1, A reaction; lane 2, G reaction; lanes 3-13, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 14, intact DNA. Putative major sites of alkylation on the DNA fragments are indicated by arrows on the sequences adjacent to each gel.

alkylation is not expected at mismatch site A since an A•T base pair is not located next to the turn unit upon conjugate binding. In addition,  $\gamma$ -DABA conjugate **6-Chl** is able to alkylate at a lower concentration than  $\alpha$ -DABA conjugate **5-Chl** (Figure IIB.4C, lanes 9).

Alkylation of the MFST3 plasmid insert was studied with polyamides **7-Chl** and **8-Chl** (Figure IIB.4D). The  $\alpha$ -DABA conjugate **7-Chl** was observed to specifically alkylate at the match site M, whereas the  $\gamma$ -DABA conjugate **8-Chl** alkylated at both the match and mismatch sites A and B. These results indicate that  $\alpha$ -DABA conjugates are alkylators with increased specificity and decreased reactivity relative to  $\gamma$ -DABA conjugates.

### Polyamide binding orientation

Previous footprinting studies indicated that glycine-linked polyamides bind DNA in extended, dimeric 2:1 complexes,<sup>16,20</sup> and the structures were confirmed by NMR studies.<sup>19,21</sup> Two possible binding modes were established (Figure IIB.5). In the “overlapped” binding mode, the two polyamides bind directly opposite one another with each heterocycle paired to another on the other strand. In the “slipped” binding mode, the polyamides integrate the



**Figure IIB.5.** Illustration of the binding modes for polyamides **3** and **7**. ‘W’ signifies A or T; ‘N’ represents any nucleotide. Sequence specificity for the slipped binding mode has been presumed based on the polyamide literature.

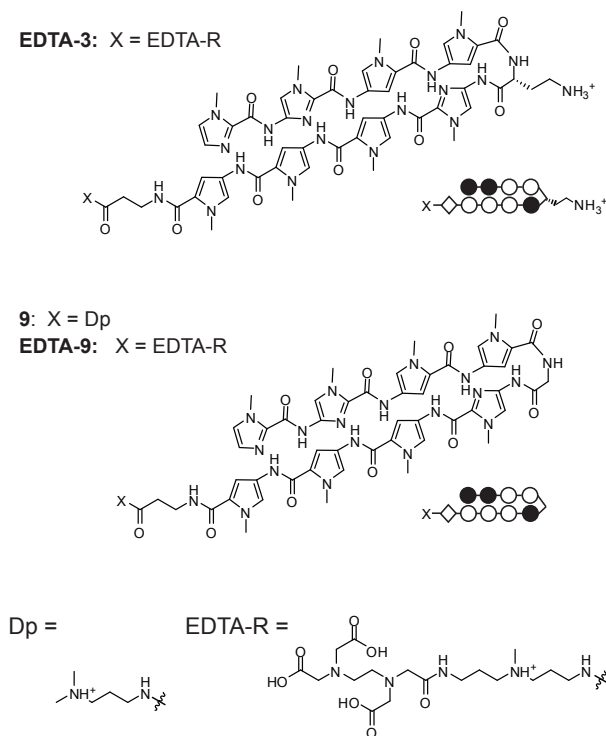


2:1 and 1:1 polyamide-DNA binding motifs at a single site.

Formally, the  $\alpha$ -DABA moiety is a substituted glycine turn, as glycine-linked polyamides also contain a single methylene group linking the heterocyclic subunits. Yet, based on their footprinting data (i.e. binding site sequence and size, and cooperativity as measured in the binding isotherms),  $\alpha$ -DABA polyamides bind in a 1:1 hairpin conformation.<sup>22</sup> Although it is well-established that  $\gamma$ -DABA polyamides bind in the forward direction,<sup>17</sup> there is still ambiguity as to the preferred binding orientation of  $\alpha$ -DABA polyamides.

We hypothesized that a polyamide with the unsubstituted glycine linkage would favor the extended conformation, whereas substitution of the turn using the  $\alpha$ -DABA moiety would force the polyamide to bind in a forward hairpin conformation. A similar phenomenon has been observed by Boger and coworkers, who studied a series of  $\alpha$ -substituted  $\beta$ -linked polyamides and found that a polyamide linked by (*R*)- $\alpha$ -methoxy- $\beta$  preferentially adopts a hairpin conformation versus the unsubstituted molecule, which binds in an extended conformation.<sup>25</sup> Interestingly, other substitutions disrupted binding affinity.

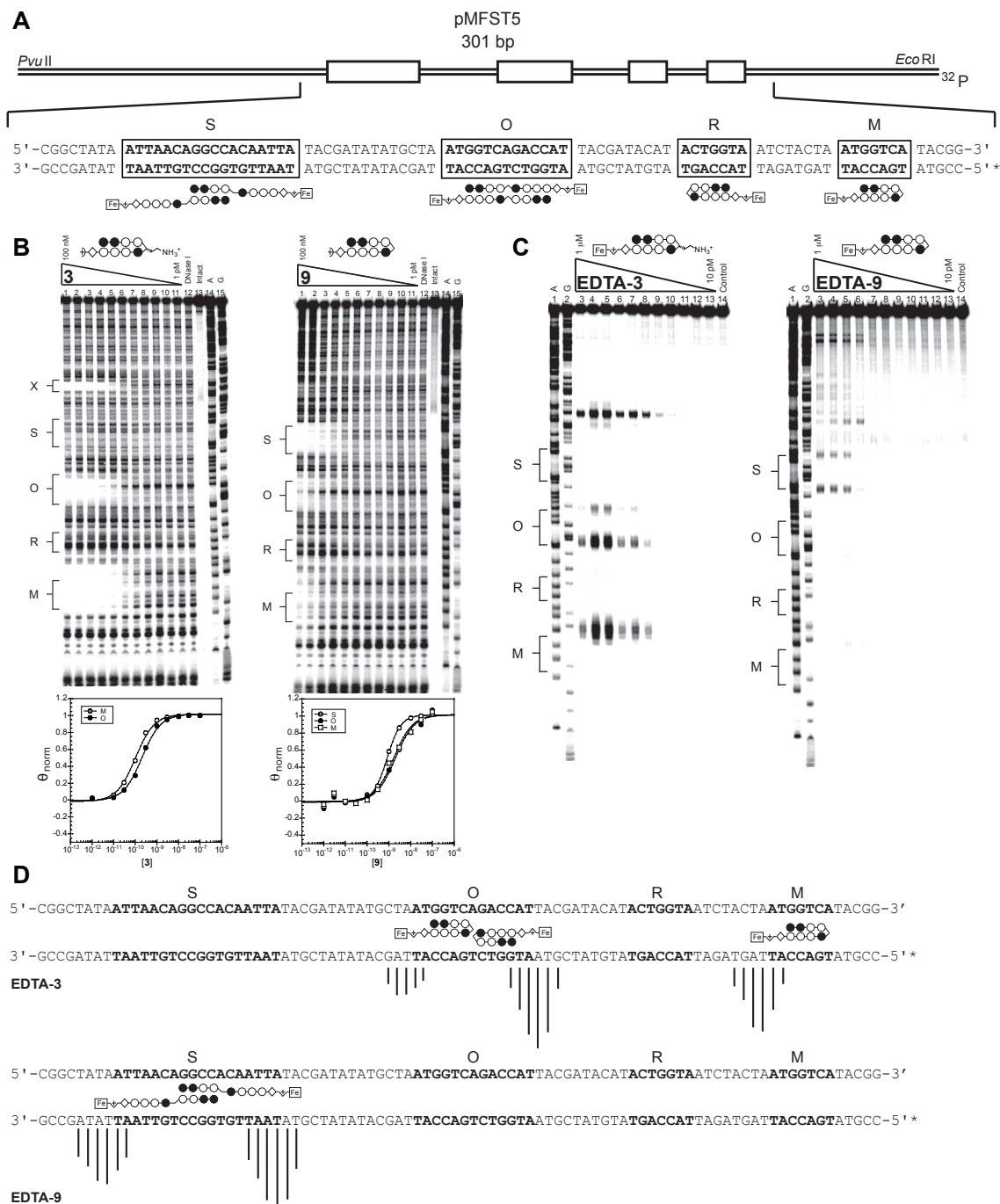
$\alpha$ -DABA polyamide **3** was chosen as the parent molecule for study. To verify its binding orientation, the EDTA conjugate **EDTA-3** was synthesized for affinity cleavage studies (Figure IIB.6). Polyamide **9** – the glycine-linked analog of **3** – was prepared, as well as its EDTA conjugate, **EDTA-9**. Four binding orientations are possible for polyamides **3** and **9**. The expected match site for the molecules binding as forward hairpins is 5'-WWGGWCW-3', whereas for reverse-binding hairpins the expected match site is 5'-WCWGGWW-3'. In the overlapped binding mode, the expected target sequence is 5'-WWGGWCWGCCWW-3'. It should be noted that the overlapped match site is inherently ambiguous because it contains two side-by-side copies of the forward hairpin match site. Binding affinity in the slipped mode is sensitive to the particular sequence of the 1:1 binding portion, and there are multiple slipped binding modes for a given polyamide



**Figure IIB.6.** Chemical and ball-and-stick structures of **EDTA-3**, and of the glycine-linked polyamide **9** and its EDTA conjugate **EDTA-9**.

core ring sequence.<sup>20</sup> Since an Im/Im pair is disfavored,<sup>26</sup> we studied the slipped orientation for **3** and **9** believed to be preferred for binding. We expect this binding model to give an 18 base pair match site with sequence 5'-WWWWWNWGGCCWNWWWW-3' where N = A, T, G, or C.<sup>20,27</sup>

Plasmid pMFST5 was prepared to study the binding preferences of the selected molecules (Figure IIB.7A). The designed insert contains four match sites for each of the possible binding modes for polyamides **3** and **9**: slipped extended binding ('S' = 5'-ATTAACAGGCCACAATTA-3'), overlapped extended binding ('O' = 5'-ATGGTCAGACCAT-3'), reverse hairpin binding ('R' = 5'-ACTGGTA-3'), and forward hairpin binding ('M' = 5'-ATGGTCA-3'). Quantitative DNase I footprinting titrations were performed with  $\alpha$ -DABA polyamide **3** and glycine polyamide **9** (Figure IIB.7B and Table IIB.2), and affinity cleavage studies were performed with  $\alpha$ -DABA conjugate **EDTA-3** and glycine conjugate **EDTA-9** (Figure IIB.7C and 7D) on the 5' <sup>32</sup>P-labeled, 301 base pair



**Figure IIB.7.** (A) Illustration of the *EcoRI*/*PvuII* restriction fragment derived from plasmid pMFST5. The designed polyamide binding sites are indicated by boxes. Binding models are illustrated below the DNA sequence by ball-and-stick structures of **EDTA-9**. (B) Quantitative DNase I footprinting titration experiments for polyamides **3** (left) and **9** (right) on the 301 base pair, 5' end-labeled PCR product of plasmid pMFST5: lanes 1-11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Site X indicates an undesigned match site inherent in the plasmid sequence. Each footprinting gel is accompanied by its respective binding isotherms (below). (C) Affinity cleavage experiments with polyamide-EDTA conjugates (left) **EDTA-3** and (right) **EDTA-9** on the 301 base pair, 5' end-labeled PCR product of plasmid pMFST5: lane 1, A reaction; lane 2, G reaction; lanes 3-13, 1  $\mu\text{M}$ , 300 nM, 100 nM, 30

(Continued from previous page)

nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, and 10 pM polyamide, respectively; lane 14, intact DNA. **(D)** Summary of affinity cleavage patterns for **EDTA-3** (top) and **EDTA-9** (bottom). Bar heights are proportional to the relative cleavage intensities at each base pair, normalized to each gel. Ball-and-stick structures illustrating their respective proposed modes of binding are shown using **EDTA-9** for clarity. Designed binding sites are indicated in bold in the DNA sequence.

PCR product of plasmid pMFST5. A dramatic difference between the two molecules was observed. The  $\alpha$ -DABA polyamide **3** binds the forward hairpin binding site M with  $K_a = 9.8 \times 10^9 \text{ M}^{-1}$ , while no binding was detected at the slipped binding site S or reverse hairpin site R. Cleavage by **EDTA-3** was observed at match site M at the locus in agreement with forward binding and is 3'-shifted, consistent with minor groove occupancy.<sup>28</sup>

Polyamide	M	O	S
<b>3</b>	$9.8 (\pm 1.3) \times 10^9$	$4.7 (\pm 1.3) \times 10^9$	--
<b>9</b>	$7.0 (\pm 1.0) \times 10^8$	$5.9 (\pm 2.1) \times 10^8$	$1.3 (\pm 0.2) \times 10^9$

**Table IIB.2.** Binding affinities ( $\text{M}^{-1}$ ) for parent polyamides **3** and **9** on plasmid pMFST5. Equilibrium association constants reported are mean values from three DNase I footprinting titration experiments. Standard deviations are shown in parentheses.

Additional binding of polyamide **3** was observed at the overlapped site O with  $K_a = 4.7 \times 10^9 \text{ M}^{-1}$ . Unequal cleavage intensities by **EDTA-3** were observed at the 3' and 5' ends of the match site. A fully anti-parallel overlapped motif would be expected to afford equal intensities. Therefore it is likely that polyamide **3** is binding as a hairpin in either of the adjacent hairpin match sites embedded in the 13 base-pair sequence O (Figure IIB.7D, top).

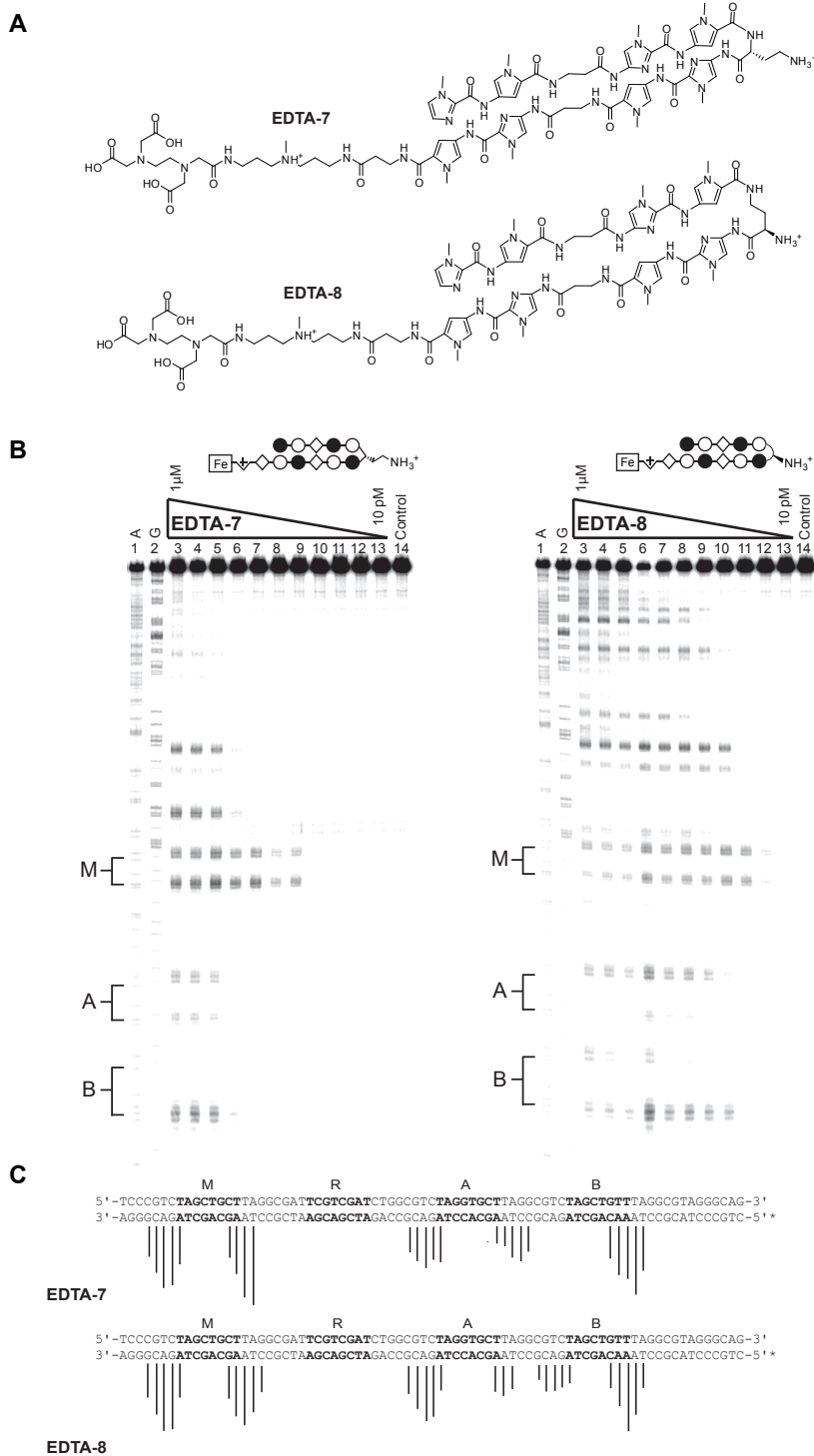
In contrast, the glycine polyamide **9** binds the slipped extended site S with  $K_a =$

$1.3 \times 10^9 \text{ M}^{-1}$ , and it binds the forward hairpin site M and overlapped site O with  $\sim 2$ -fold decreased affinities ( $K_a = 7.0 \times 10^8 \text{ M}^{-1}$  and  $K_a = 5.9 \times 10^8 \text{ M}^{-1}$ , respectively). No binding was observed at the reverse hairpin site R. Strong cleavage by **EDTA-9** was only observed at site S, probably due to the weaker binding affinities at the other sites. The cleavage pattern was 3' shifted, verifying minor groove occupancy of the ligand, and the observed pattern was consistent with dimeric, antiparallel binding in the slipped mode. There is again ambiguity regarding the binding mode at the overlapped site O. It is interesting to note that though polyamide **9** is a relatively small ligand (MW = 1196), it displays specificity for an 18 base pair sequence. It is likely that the binding affinity of **9** in the slipped binding mode could be enhanced by modifying the sequence of the W tracts in the 1:1 binding portions of the molecules to a DNA sequence that favors a narrower minor groove.<sup>20</sup>

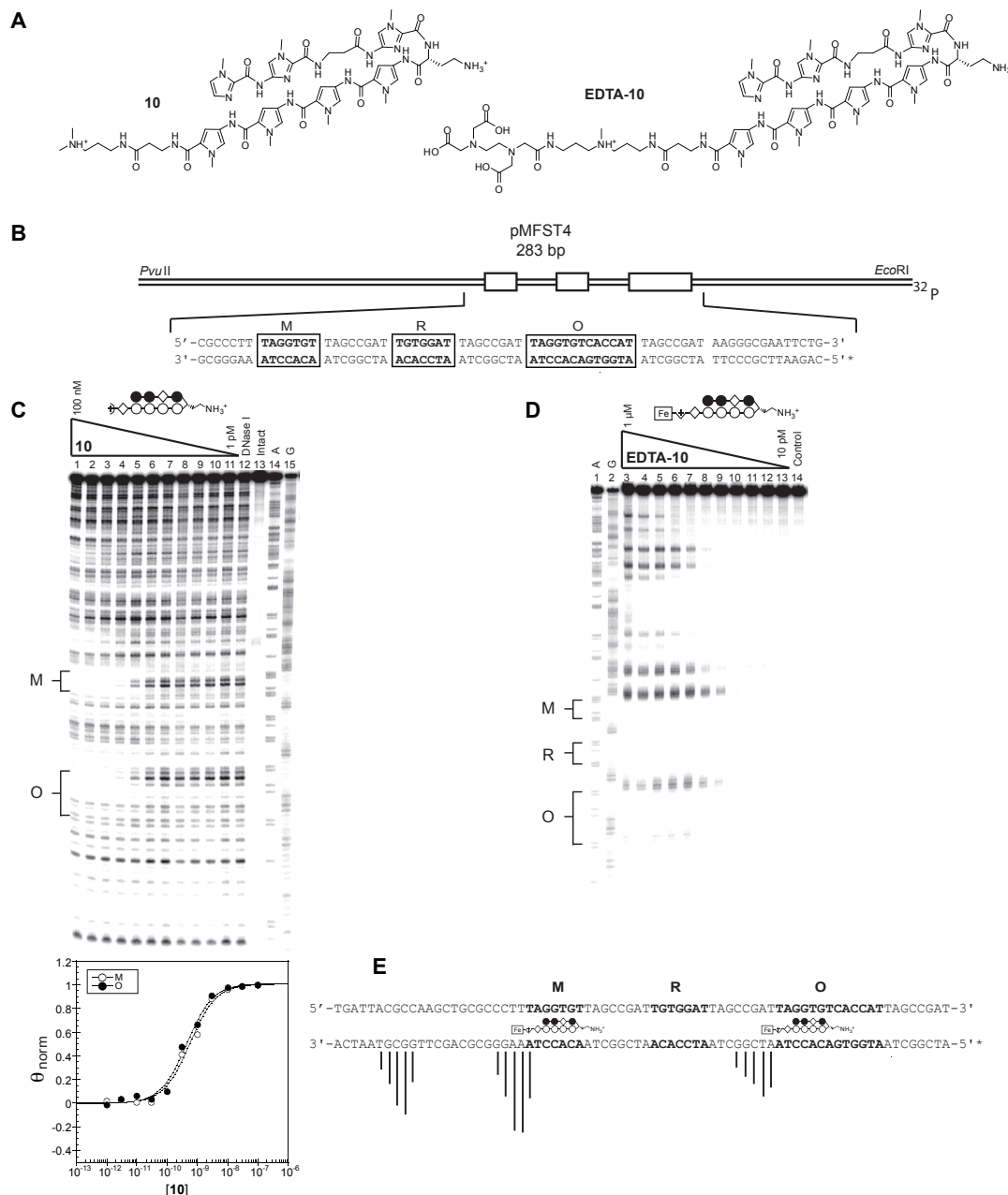
These data show a clear distinction in preferred binding modes between  $\alpha$ -DABA polyamide **3** and glycine polyamide **9**. The  $\alpha$ -DABA polyamide appears to bind exclusively in a forward hairpin conformation, while the glycine polyamide seems to exhibit mixed binding motifs, preferring the extended binding mode but also appearing to be capable of binding as a forward hairpin. Thus, substitution of the glycine turn 'locks' the polyamide into a hairpin conformation.

Additional affinity cleavage studies were performed on the EDTA conjugates of polyamides **7** and **8**, **7-EDTA** and **8-EDTA**, respectively (Figure IIB.8A), and **EDTA-10**, the EDTA conjugate of polyamide **10**, which has been previously described (Figure IIB.9A).<sup>22</sup> Affinity cleavage experiments for **EDTA-7** and **EDTA-8** were performed on pMFST3 (Figure IIB.2C and 8B). Cleavage was observed at both 5' and 3' sites flanking the polyamides (Figure IIB.8C); this is likely due to the nature of the targeted DNA sequence, which allows for the molecules to bind in a forward orientation, 5'→3' with respect to either the top or bottom strands.

Plasmid pMFST4 (Figure IIB.9B) was designed to study the binding mode of polyamide **10**, and was designed with both a forward and reverse match site ('M' = 5'-



**Figure IIB.8.** (A) Chemical structures for polyamide-EDTA conjugates **EDTA-7** and **EDTA-8**. (B) Affinity cleavage experiments with polyamide-EDTA conjugates **EDTA-7** (left) and **EDTA-8** (right) on the 288 base pair, 5' end-labeled PCR product of plasmid pMFST3: lane 1, A reaction; lane 2, G reaction; lanes 3-13, 1  $\mu$ M, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, and 10 pM polyamide, respectively; lane 14, intact DNA. (C) Summary of affinity cleavage patterns for (top) **EDTA-7** (top) and **EDTA-8** (bottom). Bar heights are proportional to the relative cleavage intensities at each base pair. The designed polyamide binding sites are indicated in bold.



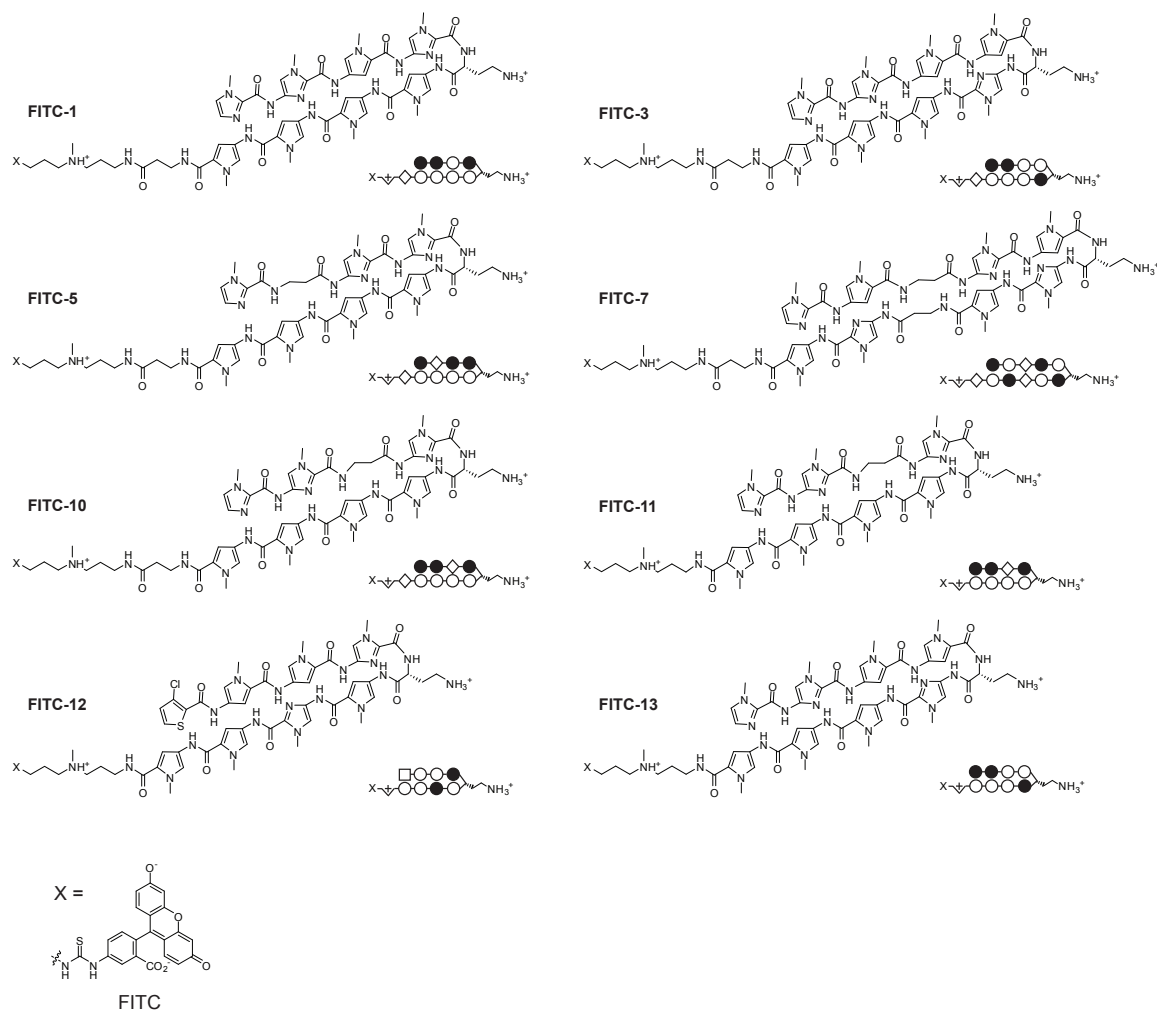
**Figure IIB.9.** (A) Chemical structures for polyamide **10** and its EDTA conjugate **EDTA-10**. (B) Illustration of the *EcoRI*/*PvuII* restriction fragment derived from plasmid pMFST4. The designed polyamide binding sites are indicated by boxes. (C) Quantitative DNase I footprinting titration experiments for polyamide **10** on the 283 base pair, 5' end-labeled PCR product of plasmid pMFST4: lanes 1-11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Binding isotherms (below). Binding affinities for sites M and O are  $9.8 (\pm 0.5) \times 10^8 \text{ M}^{-1}$  and  $1.1 (\pm 0.6) \times 10^9 \text{ M}^{-1}$ , respectively. (D) Affinity cleavage experiment with **EDTA-10** on the 283 base pair, 5' end-labeled PCR product of plasmid pMFST4: lane 1, A reaction; lane 2, G reaction; lanes 3-13, 1  $\mu\text{M}$ , 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, and 10 pM polyamide, respectively; lane 14, intact DNA. (E) Summary of affinity cleavage patterns for **EDTA-10**. Bar heights are proportional to the relative cleavage intensities at each base pair. Ball-and-stick structures illustrating the proposed modes of binding are shown. Designed binding sites are indicated in bold in the DNA sequence.

TAGGTGT-3' and 'R' = 5'-TGTGGAT-3', respectively), and an overlapped binding site ('O' = 5'-TAGGTGTCACCAT-3'). DNase I footprint titrations were performed using the parent polyamide **10** (Figure IIB.9C). Cleavage patterns observed are consistent with polyamide binding as a hairpin in the forward orientation (Figure IIB.9D and 9E).

## Conclusion

Herein we present a characterization of  $\alpha$ -DABA polyamides. We have confirmed their ability to bind DNA and established their binding affinities to be decreased 2-fold to 147-fold relative to  $\gamma$ -DABA polyamides. The magnitude of this reduced affinity appears to be sensitive to the presence and positioning of an internal  $\beta$  residue in the polyamide structure. Remarkably, despite the reduced affinity,  $\alpha$ -DABA polyamide-chlorambucil conjugates alkylate DNA with greater specificity and lower reactivity than  $\gamma$ -DABA conjugates. Importantly, we have demonstrated that polyamides linked by the (*R*)- $\alpha$ -diaminobutyric acid moiety bind DNA in a forward hairpin binding mode. Polyamides linked by a glycine amino acid prefer to bind DNA in an extended binding mode, indicating that it is the substitution of the glycine turn unit that locks the  $\alpha$ -DABA polyamide into the hairpin conformation. While it has been previously shown that  $\alpha$ -DABA polyamide-chlorambucil conjugates are efficacious in cellular and small animal models,<sup>13,22,29</sup> cellular uptake experiments of polyamide-fluorescein conjugates displayed limited nuclear uptake properties (Figure IIB.10, Table IIB.3), possibly limiting the general use of  $\alpha$ -DABA polyamides *in vivo*.





**Figure IIB.10.** Chemical and ball-and-stick structures for polyamide-fluorescein conjugates prepared for studying the cellular uptake properties of  $\alpha$ -DABA polyamides.



## Polyamide synthesis

Polyamides were synthesized on Boc- $\beta$ -Ala-Pam resin following manual solid-phase methods using Boc-protected amino acids as previously reported.<sup>23</sup> Boc-Py-OBt (OBt = benzotriazol-1-yloxy), Boc- $\beta$ -Im-OH, Im-CCl<sub>3</sub>, Boc-Py-Py-OBt, Boc-Py-Im-OH, Boc-Im-OH, and Im-Im-OH were used as building blocks. The  $\alpha$ -DABA polyamides **1**, **3**, **5**, **7**, **10**, **1-Chl**, **3-Chl**, **5-Chl**, **7-Chl**, **EDTA-3**, **EDTA-7**, and **EDTA-10** were synthesized using Boc-D-Dab(Fmoc)-OH as the turn unit. The  $\gamma$ -DABA polyamides **2**, **4**, **6**, **8**, **2-Chl**, **4-Chl**, **6-Chl**, **8-Chl**, and **8-EDTA** were synthesized using Fmoc-D-Dab(Boc)-OH as the turn unit. The glycine-linked polyamides **9** and **EDTA-9** were synthesized using Boc-Gly-OH as the turn unit.

Boc deprotections were conducted at room temperature for 25 min using 80% TFA/DCM. Carboxylic acids (3 eq) were activated with DIEA (5 eq) and PyBOP (2.5 eq) for 30-45 min at 37 °C prior to coupling. Couplings were allowed to proceed in NMP for 2 h at room temperature. Polyamides were cleaved from resin with either 3-(dimethylamino)-propylamine (Dp) or 3,3'-diamino-N-methyl-dipropylamine (Da) neat at 55 °C overnight. Products were purified by preparatory reverse-phase high performance liquid chromatography (prep HPLC) and characterized by analytical HPLC, UV-visible spectroscopy, and matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry.

Polyamide-chlorambucil conjugates were generated by activating the chlorambucil carboxylic acid (4 eq) with PyBOP (4 eq) and DIEA (12 eq) for 30 min at room temperature followed by addition of the polyamide. The reaction was allowed to proceed at room temperature for 1-2 h. Products were purified by reverse-phase prep HPLC, immediately lyophilized, and characterized as above.

Polyamide-EDTA conjugates were synthesized by slow addition of a 50% DMF/DIEA polyamide solution to a solution of EDTA dianhydride (20 eq) dissolved in a 1:1:2 DMSO:DMF:DIEA mixture, vigorously shaking at 55 °C. The reaction was allowed to

proceed an additional 30 min, followed by hydrolysis of the remaining EDTA dianhydride with 0.1 N NaOH for 10 min. The reaction was diluted and purified by reverse-phase prep HPLC. Subsequent Boc deprotection was accomplished by dissolving the dried EDTA conjugate in a minimum volume of neat TFA doped with triethylsilane. The reaction was allowed to proceed for 10-30 min at room temperature, then immediately diluted with water and purified by reverse-phase prep HPLC.

Polyamide-fluorescein conjugates were synthesized by dissolving the cleaved polyamide in DMF, followed by addition of DIEA (50 eq). A solution of 5-fluorescein isothiocyanate (Molecular Probes) in DMF (4 eq) was added to the polyamide. The reaction was allowed to proceed for 40 min at room temperature. The polyamide was then precipitated by addition of chilled ether (~90 mL). The dried polyamide precipitate was then dissolved in a minimal volume of trifluoroacetic acid with triethylsilane solution. The reaction was allowed to shake at room temperature for 20 min, followed by dilution with water and acetonitrile and immediate purification by prep HPLC.

- 1: ImImPyIm-(R)<sup>H2N</sup> $\alpha$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS  $C_{57}H_{72}N_{23}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1238.58; found: 1238.82
- 2: ImImPyIm-(R)<sup>H2N</sup> $\gamma$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS  $C_{57}H_{72}N_{23}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1238.58; found: 1238.59
- 3: ImImPyPy-(R)<sup>H2N</sup> $\alpha$ -ImPyPyPy- $\beta$ Dp** MALDI-TOF MS  $C_{57}H_{72}N_{23}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1238.58; found: 1238.90
- 4: ImImPyPy-(R)<sup>H2N</sup> $\gamma$ -ImPyPyPy- $\beta$ Dp** MALDI-TOF MS  $C_{57}H_{72}N_{23}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1238.58; found: 1238.65
- 5: Im $\beta$ ImIm-(R)<sup>H2N</sup> $\alpha$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS  $C_{54}H_{71}N_{22}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1188.28; found: 1188.5
- 6: Im $\beta$ ImIm-(R)<sup>H2N</sup> $\gamma$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS  $C_{54}H_{71}N_{22}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1188.28; found: 1188.86

**7: ImPyβImPy-(R)<sup>H2N</sup>α-ImPyβImPy-βDp** MALDI-TOF MS  $C_{62}H_{81}N_{26}O_{12}^{+}$  calculated [M+H]<sup>+</sup>: 1382.47; found: 1382.15

**8: ImPyβImPy-(R)<sup>H2N</sup>γ-ImPyβImPy-βDp** MALDI-TOF MS  $C_{62}H_{81}N_{26}O_{12}^{+}$  calculated [M+H]<sup>+</sup>: 1382.47; found: 1382.58

**9: ImImPyPy-Gly-ImPyPyPy-βDp** MALDI-TOF MS  $C_{55}H_{67}N_{22}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1195.54; found: 1195.43

**10: ImImβIm-(R)<sup>H2N</sup>α-PyPyPyPy-βDp** MALDI-TOF  $C_{54}H_{71}N_{22}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1188.28, found: 1188.02

**1-Chl: ImImPyIm-(R)<sup>Chl</sup>α-PyPyPyPy-βDp** MALDI-TOF MS  $C_{71}H_{89}Cl_2N_{24}O_{11}^{+}$  calculated [M+H]<sup>+</sup>: 1523.65; found: 1523.66

**2-Chl: ImImPyIm-(R)<sup>Chl</sup>γ-PyPyPyPy-βDp** MALDI-TOF MS  $C_{71}H_{89}Cl_2N_{24}O_{11}^{+}$  calculated [M+H]<sup>+</sup>: 1523.65; found: 1523.70

**3-Chl: ImImPyPy-(R)<sup>Chl</sup>α-ImPyPyPy-βDp** MALDI-TOF MS  $C_{71}H_{89}Cl_2N_{24}O_{11}^{+}$  calculated [M+H]<sup>+</sup>: 1523.65; found: 1523.88

**4-Chl: ImImPyPy-(R)<sup>Chl</sup>γ-ImPyPyPy-βDp** MALDI-TOF MS  $C_{71}H_{89}Cl_2N_{24}O_{11}^{+}$  calculated [M+H]<sup>+</sup>: 1523.65; found: 1523.57

**5-Chl: ImβImIm-(R)<sup>Chl</sup>α-PyPyPyPy-βDp** MALDI-TOF MS  $C_{68}H_{88}Cl_2N_{23}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1474.48, found: 1474.95

**6-Chl: ImβImIm-(R)<sup>Chl</sup>γ-PyPyPyPy-βDp** MALDI-TOF MS  $C_{68}H_{88}Cl_2N_{23}O_{10}^{+}$  calculated [M+H]<sup>+</sup>: 1474.48, found: 1474.26

**7-Chl: ImPyβImPy-(R)<sup>Chl</sup>α-ImPyβImPy-βDp** MALDI-TOF MS  $C_{76}H_{98}Cl_2N_{27}O_{13}^{+}$  calculated [M+H]<sup>+</sup>: 1668.67, found: 1668.37

**8-Chl: ImPyβImPy-(R)<sup>Chl</sup>γ-ImPyβImPy-βDp** MALDI-TOF MS  $C_{76}H_{98}Cl_2N_{27}O_{13}^{+}$  calculated [M+H]<sup>+</sup>: 1668.67, found: 1668.48

**EDTA-3: ImImPyPy-(R)<sup>H2N</sup>α-ImPyPyPy-βDa-EDTA** MALDI-TOF MS  $C_{69}H_{91}N_{26}O_{17}^{+}$  calculated [M+H]<sup>+</sup>: 1555.71; found: 1555.73

**EDTA-7: ImPyβImPy-(R)<sup>H2N</sup>α-ImPyβImPy-βDa-EDTA** MALDI-TOF MS

$C_{74}H_{100}N_{29}O_{19}^{+}$  calculated  $[M+H]^{+}$ : 1698.77, found: 1698.81

**EDTA-8: ImPy $\beta$ ImPy-(R)<sup>H<sup>2</sup>N</sup> $\gamma$ -ImPy $\beta$ ImPy- $\beta$ Da-EDTA** MALDI-TOF MS

$C_{76}H_{98}Cl_2N_{27}O_{13}^{+}$  calculated  $[M+H]^{+}$ : 1698.77, found: 1698.58

**EDTA-9: ImImPyPy-Gly-ImPyPyPy- $\beta$ Da-EDTA** MALDI-TOF MS  $C_{67}H_{86}N_{25}O_{17}^{+}$

calculated  $[M+H]^{+}$ : 1512.66; found: 1512.54

**EDTA-10: ImIm $\beta$ Im-(R)<sup>H<sup>2</sup>N</sup> $\alpha$ -PyPyPyPy- $\beta$ Da-EDTA** MALDI-TOF MS  $C_{66}H_{90}N_{25}O_{17}^{+}$

calculated  $[M+H]^{+}$ : 1504.70; found: 1504.84

### Construction of plasmids

Plasmids were constructed by annealing oligonucleotide pairs as follows: for pMFST – 5'-AGCTGTACAATCATTAGTGGTTACAATCATATGGTCATACAATC-ATTAGTCGTTACAATCATTAGCACACACAATCATC-3' and 5'-GATCGATGATTGTGTGTGCTAATGATTGTAACGACTAATGATTGTATGACCATATGATTGTAA-CCACTAATGATTGTAC-3'; for pMFST5 – 5'-AGCTGCGGCTATAATTAACAGGC-CACAATTATACGATATATGCTAATGGTCAGACCATTACGATACATACTGGTA-ATCTACTAATGGTCATACGGC-3' and 5'-GATCGCCGTATGACCATTAGTAGAT-TACCAGTATGTATCGTAATGGTCTGACCATTAGCATATATCGTATAATTGTG-GCCTGTTAATTATAGCCGC-3'.

Annealed nucleotides were ligated into the *Bam*HI/*Hind*III restriction fragment of pUC19 using T4 DNA ligase, and the plasmid was transformed into *Escherichia coli* JM109 competent cells. Ampicillin-resistant white colonies were selected from 25 mL Luria-Bertani agar plates containing 50 mg/mL ampicillin treated with XGAL and isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) solutions and grown overnight at 37 °C. Cells were harvested the following day, and purification of the plasmids were performed with a Wizard Plus Midiprep DNA purification kit (Promega). DNA sequencing of the plasmid insert was performed by the sequence analysis facility at the California Institute of Technology. Plasmids pMFST3 and pMFST4 were prepared similarly with oligonucleotides of the

indicated sequences.

### **Preparation of 5' <sup>32</sup>P-end-labeled DNA**

The primer 5'-GAATTCGAGCTCGGTACCCGGG-3' was <sup>32</sup>P-labeled at the 5' end and used with the primer 3'-CAGCCCTTTGGACAGCACGGTC-5' to PCR amplify plasmids as described previously.<sup>28</sup>

### **DNase footprinting titrations**

Polyamide equilibrations and DNase I footprinting titrations were conducted on the 5' end-labeled PCR product of the appropriate plasmid according to standard protocols.<sup>28</sup> DNA was incubated with polyamide or water (controls) for 12-16 h at room temperature prior to reaction. Determination of equilibrium association constants were performed as previously described.<sup>28</sup> Data points were fitted to a Langmuir binding isotherm by using the modified Hill equation with  $n = 1$ , except for the quantitation of glycine polyamide **9**, where  $n$  was optimized to the generated curve for the slipped site S and the overlapped site O.

### **Thermal cleavage assays**

Thermal cleavage assay experiments were conducted on the 5' end-labeled PCR products of the designated plasmids as described previously.<sup>11</sup> DNA was incubated with polyamide conjugates or water (control) for 24 h at 37 °C prior to work-up.

### **Affinity cleavage assays**

Affinity cleavage experiments were conducted on the 5' end-labeled PCR product of plasmids pMFST3, pMFST4, and pMFST5 as previously described.<sup>28</sup> DNA was incubated with polyamide conjugates or water (control) for 13 h at room temperature prior to addition of ferrous ammonium sulfate.

## Acknowledgements

The authors thank the National Institutes of Health for research support and for a pre-doctoral NRSA training grant to M.E.F. Mass spectrometry analyses were performed in the Mass Spectrometry Laboratory of the Division of Chemistry and Chemical Engineering at the California Institute of Technology.

## References

1. Wemmer, D. E.; Dervan, P. B. *Curr. Opin. Struct. Biol.* **1997**. *7*, 355-361.
2. Dervan, P. B.; Poulin-Kerstien, A. T.; Fechter, E. J.; Edelson, B. S. *Top. Curr. Chem.* **2005**. *253*, 1-31.
3. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003**. *100*, 12063-12068.
4. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004**. *32*, 2802-2818.
5. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2004**. *101*, 16768-16773.
6. Burnett, R.; Melander, C.; Puckett, J. W.; Son, L. S.; Wells, R. D.; Dervan, P. B.; Gottesfeld, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2006**. *103*, 11497-11502.
7. Kageyama, Y.; Sugiyama, H.; Ayame, H.; Iwai, A.; Fujii, Y.; Huang, L. E.; Kizaka-Kondoh, S.; Hiraoka, M.; Kihara, K. *Acta Oncol.* **2006**. *45*, 317-324.
8. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**. *35*, 363-370.
9. Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2007**. *104*, 10418-10423.
10. Bando, T.; Sugiyama, H. *Accounts Chem. Res.* **2006**. *39*, 935-944.
11. Wurtz, N. R.; Dervan, P. B. *Chem. Biol.* **2000**. *7*, 153-161.
12. Wang, Y. D.; Dziegielewski, J.; Wurtz, N. R.; Dziegielewska, B.; Dervan, P. B.; Beerman, T. A. *Nucleic Acids Res.* **2003**. *31*, 1208-1215.



13. Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B. S.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M. *Chem. Biol.* **2004**. *11*, 1583-1594.
14. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**. *120*, 6219-6226.
15. Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1992**. *89*, 7586-7590.
16. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**. *116*, 7983-7988.
17. White, S.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**. *119*, 8756-8765.
18. Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**. *120*, 1382-1391.
19. Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Nat. Struct. Biol.* **1996**. *3*, 321-324.
20. Trauger, J. W.; Baird, E. E.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**. *118*, 6160-6166.
21. de Clairac, R. P. L.; Seel, C. J.; Geierstanger, B. H.; Mrksich, M.; Baird, E. E.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1999**. *121*, 2956-2964.
22. Tsai, S. M.; Farkas, M. E.; Chou, C. J.; Gottesfeld, J. M.; Dervan, P. B. *Nucleic Acids Res.* **2007**. *35*, 307-316.
23. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**. *118*, 6141-6146.
24. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B.; Puckett, J. W.; Wang, C. C.; Dervan, P. B. *Tetrahedron*. **2007**. *63*, 6146-6151.
25. Woods, C. R.; Ishii, T.; Wu, B.; Bair, K. W.; Boger, D. L. *J. Am. Chem. Soc.* **2002**. *124*, 2148-2152.
26. White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**. *4*, 569-578.
27. Urbach, A. R.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2001**. *98*, 4343-4348.
28. Trauger, J. W.; Dervan, P. B. *Meth. Enzymol.* **2001**. *340*, 450-466.
29. Alvarez, D.; Chou, C. J.; Latella, L.; Zeitlin, S. G.; Ku, S.; Puri, P. L.; Dervan, P. B.; Gottesfeld, J. M. *Cell Cycle*. **2006**. *5*, 1537-1548.

## Chapter IIC

### **Small Molecules Targeting Histone H4 as Potential Therapeutics for Chronic Myelogenous Leukemia**

*The text of this chapter was taken in part from a manuscript co-authored with Sherry M. Tsai,<sup>\*</sup> Peter B. Dervan,<sup>\*</sup> C. James Chou,<sup>‡</sup> David Alvarez,<sup>‡</sup> and Joel M. Gottesfeld<sup>‡</sup> ( <sup>\*</sup> California Institute of Technology and <sup>‡</sup> The Scripps Research Institute)*

(Chou, C.J.; Farkas, M.E.; Tsai, S.M.; Alvarez, D.; Gottesfeld, J.M.; Dervan, P.B. *Mol. Cancer Ther.* **2008**, 7, 769-778)

## Abstract

We recently identified a polyamide-chlorambucil conjugate, **1R-Chl**, that alkylates and down-regulates transcription of the human histone H4c gene, and inhibits the growth of several cancer cell lines *in vitro* and in a murine SW620 xenograft model without apparent animal toxicity. In this study, we analyzed the effects of **1R-Chl** in the chronic myelogenous leukemia cell line K562 and identified another polyamide conjugate, **6R-Chl**, which targets H4 genes and elicits a similar cellular response. Other polyamide conjugates that do not target the H4 gene do not elicit this response. Gene chip expression analysis has shown that only 156 genes ( $p = 0.00125$ ) are affected by **1R-Chl** in K562 cells, and has identified three genes of interest encoding histone H4: H4c and H4k/j. In a murine model, both **1R-Chl** and **6R-Chl** were found to be highly effective in blocking K562 xenograft growth with high dose tolerance. Unlike with conventional and distamycin-based alkylators, little or no cyto- and animal toxicities were observed in mg/kg dosage ranges. These results suggest that these polyamide alkylators may be a viable treatment alternative for chronic myelogenous leukemia.

## Introduction

Therapies targeting specific genes or gene products are a major aim of modern cancer biology and intervention.<sup>1-4</sup> Notable recent successes in the field include both monoclonal antibodies<sup>5,6</sup> and small molecules that target enzymes and receptors that are over-expressed or mutated in various cancers.<sup>3,7,8</sup> Notwithstanding these important developments, DNA alkylating agents remain the most common drugs for the treatment of several solid and hematological malignancies.<sup>9-11</sup> Myelotoxicity and concomitant cytotoxicity due to limited DNA sequence selectivity is often the dose-limiting factor for use of these compounds in humans. Alkylating agents based on the minor groove binder distamycin A, including Tallimustine and Brostallicin, have shown improvements over non-conjugated DNA alkylating agents in terms of affinity and specificity at the nucleotide level.<sup>12-15</sup> However, these compounds possess only limited specificity,<sup>16</sup> and myelotoxicity is still the dose-limiting factor in establishing an effective chemotherapy.<sup>12,13</sup>

Pyrrole-imidazole (Py-Im) polyamides are a class of sequence-specific DNA-binding small molecules that have been shown to have high binding specificity and affinity, with some molecules even exceeding the binding affinities of transcription factors.<sup>17</sup> Many polyamides are cell permeable and readily localize to the nuclei of cultured cells.<sup>18-21</sup> Polyamides are effective inhibitors of RNA transcription when they disrupt essential protein-DNA interactions at promoter and enhancer elements,<sup>17,22-25</sup> and they may also modulate gene expression by modifying chromatin structure.<sup>26-28</sup> Linked with DNA-alkylating agents, such as chlorambucil (Chl)<sup>29</sup> or CC-1065/CBI derivatives,<sup>30,31</sup> polyamide conjugates react covalently with DNA at specific sites and inhibit transcription by stalling RNA polymerase during elongation.

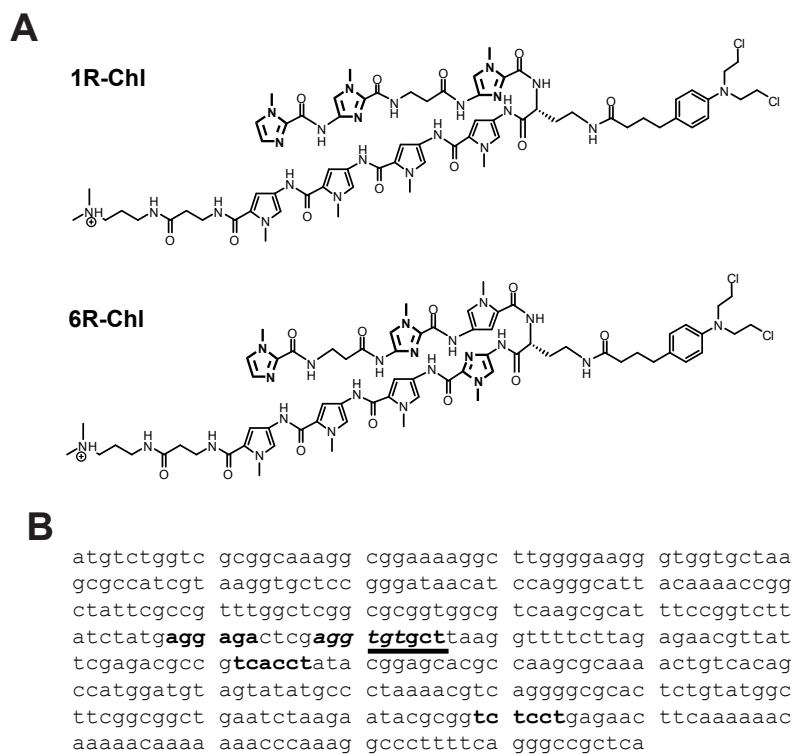
Recently,  $\alpha$ -diaminobutyric acid-linked hairpin polyamide-Chl conjugates have been shown to bind and alkylate DNA sequences both *in vitro* and in cell culture models, with good sequence specificity.<sup>32-34</sup> Initial screens of various hairpin polyamide-Chl conjugates have shown the polyamide **1R-Chl** (Figure IIC.1, top) to be an inhibitor of

cell proliferation in various cancer cell lines with no apparent cytotoxicity and little or no murine animal toxicity.<sup>32,33,35</sup> This molecule binds within the coding region of the histone H4c gene both *in vitro* and in SW620 human colon carcinoma cells, and down-regulates H4c transcription. Polyamides with similar pyrrole and imidazole compositions targeted to different DNA sequences failed to alkylate the coding region of the histone H4c gene and were found to be inactive in both cell culture and a SW620 xenograft cancer model.<sup>32</sup>

Studies suggest a two-hit mechanism for the observed cellular effects of **1R-Chl**: down-regulation of histone gene transcription causes nucleosome depletion, followed by widespread alkylation of open chromatin, which elicits cell cycle arrest through the DNA repair pathway.<sup>35</sup> While our initial results point to the histone H4c gene as the major target of **1R-Chl**, microarray studies in the SW620 cancer cell line indicate that the mRNA levels of several other genes are also affected.<sup>32</sup> Thus, down-regulation of other genes may be involved in the cellular response to **1R-Chl**.

In the present study we extend our analysis to the well-established chronic myelogenous leukemia (CML) cell line K562. If histone H4 genes are the primary targets of **1R-Chl** that lead to a block in cancer cell proliferation,<sup>35</sup> other polyamide-Chl conjugates targeting H4 genes would be predicted to elicit the same cellular response. We describe the synthesis and characterization of a small library of constitutional isomers of **1R-Chl**. These molecules have the same chemical composition as **1R-Chl** but would be expected to bind different DNA sequences. One conjugate, **6R-Chl** (Figure IIC.1A, bottom), which targets sites adjacent to and overlapping the binding site for **1R-Chl** (Figure IIC.1B) in the H4c gene was found to have biological properties similar to **1R-Chl** in both K562 cell culture and in a mouse xenograft model established with K562 cells. Other polyamide-Chl alkylators that did not bind within the H4c gene or down-regulate histone H4 expression had no effect on cell proliferation. Microarray analysis in K562 cells reveals that the histone H4 genes H4c and H4j/k are down-regulated by **1R-Chl** treatment. Transcripts for the H4k and H4j genes cannot be distinguished due to similarity in sequence. In addition,

we explored the pharmacokinetic properties of **1R-Chl**, the results of which point to this class of molecules as potential human cancer therapeutics.






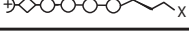


**Figure IIC.1.** DNA sequence of the coding region of the human histone H4c gene and chemical structures of **1R-Chl** and **6R-Chl**. (A) Chemical structures of **1R-Chl** (top) and **6R-Chl** (bottom), which target the DNA sequences 5'-WGGWGW-3' and 5'-WGGWCW-3', respectively (where W = A or T). Imidazole rings are shown in bold. (B) DNA sequence of the coding region of the human histone H4c gene. Potential binding sites for **1R-Chl** and **6R-Chl** are in bold. The alkylation sites of **1R-Chl** and **6R-Chl** as verified by LM-PCR are italic-bold and underlined, respectively.

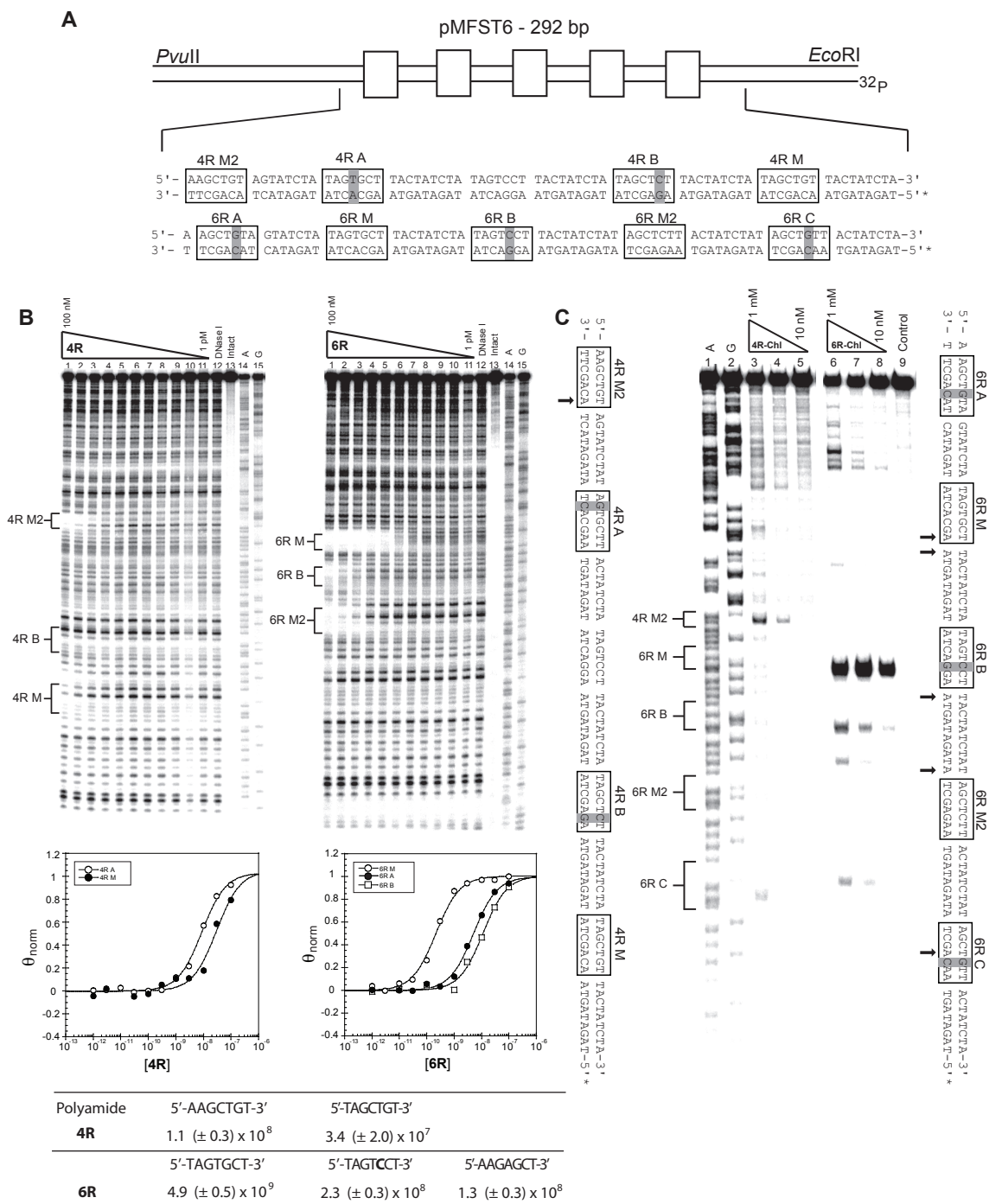
## Results and Discussion

### Inhibition of K562 cell proliferation by polyamides targeting histone H4 genes

To investigate the effects of Py-Im polyamide-Chl conjugates on K562 cells, we synthesized a small library of constitutional isomers of **1R-Chl**, along with its stereoisomer, **1S-Chl** (Table IIC.1). All molecules targeted the general DNA sequence 5'-WGNNW-3' (where W = A or T; N = A, C, G, or T); the binding site for **1R-Chl** is 5'-WGGWGW-3'. Binding affinities for non-alkylating parent molecules were determined via quantitative DNase I footprinting on plasmid inserts containing matched sites. It has been previously shown that attachment of the Chl moiety does not appear to affect DNA binding properties.<sup>29</sup> Affinities for **1R**, **3R**, **5R**, and **1S** have been previously reported (Table IIC.1);<sup>33-35</sup> polyamides **4R** and **6R** were analyzed on plasmid insert pMFST6 (Figure IIC.2A and 2B). All polyamides were found to bind their expected match sites.

Polyamide	Sequence	Match Site (Plasmid)	$K_a$ Match Site ( $M^{-1}$ ) $X = NH_3^+$	Cell Enlargement	Growth Arrest
<b>1R</b> 	5'-WGGWGW-3'	5'-TAGGTGT-3' (pMFST2)	$6.7(\pm 0.1) \times 10^9$	yes	yes (>50%)
<b>3R</b> 	5'-WGGWCW-3'	5'-TAGGTCT-3' (pJG)	$7.4(\pm 0.6) \times 10^8$	no	no
<b>4R</b> 	5'-WGCWGW-3'	5'-AAGCTGT-3' (pMFST6)	$1.1 (\pm 0.3) \times 10^8$	no	no
<b>5R</b> 	5'-WGWGGW-3'	5'-TAGTGGT-3' (pMFST)	$8.0 (\pm 0.6) \times 10^8$	no	slight (~25%)
<b>6R</b> 	5'-WGWGCW-3'	5'-TAGTGCT-3' (pMFST6)	$4.9 (\pm 0.5) \times 10^9$	yes	yes (>70%)
<b>1S</b> 	5'-WGWGGW-3'	5'-TAGGTGT-3' (pMFST2)	$4.8 (\pm 0.9) \times 10^8$	no	no

**Table IIC.1.** Analysis of a polyamide library. Shown in the table are the following: ball-and-stick model for the structure of each polyamide (closed circles are imidazoles, open circles are pyrroles, diamonds are  $\beta$ -alanines, semicircle with positive sign is dimethylaminopropylamine, and X indicates a free amino group in the parent polyamide or Chl in the conjugates); targeted match binding sites for the polyamides, where 'W' = A or T, (names of plasmid constructs on which footprinting and alkylation assays were performed are shown in parentheses); equilibrium association constants of the parent polyamides as determined by quantitative DNase I footprinting (standard deviations are in parentheses); effects of polyamide-Chl conjugates on cell morphology (enlargement) and growth.



**Figure IIC.2.** Polyamide analysis on pMFST6. (A) Illustration of the *EcoRI*/*PvuII* restriction fragments derived from plasmid pMFST6. The designed polyamide binding sites are indicated by boxes. Single base pair mismatches are indicated by shaded regions. (B) Quantitative DNaseI footprinting titration experiments for polyamides **4R** (left) and **6R** (right) on the 5' end-labeled PCR product of plasmid pMFST6: lanes 1-11, 100 nM, 1 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by its respective binding isotherms (below). Binding affinities ( $M^{-1}$ ) for polyamides **4R** and **6R**, with equilibrium association constants reported as mean values from three DNase I footprinting

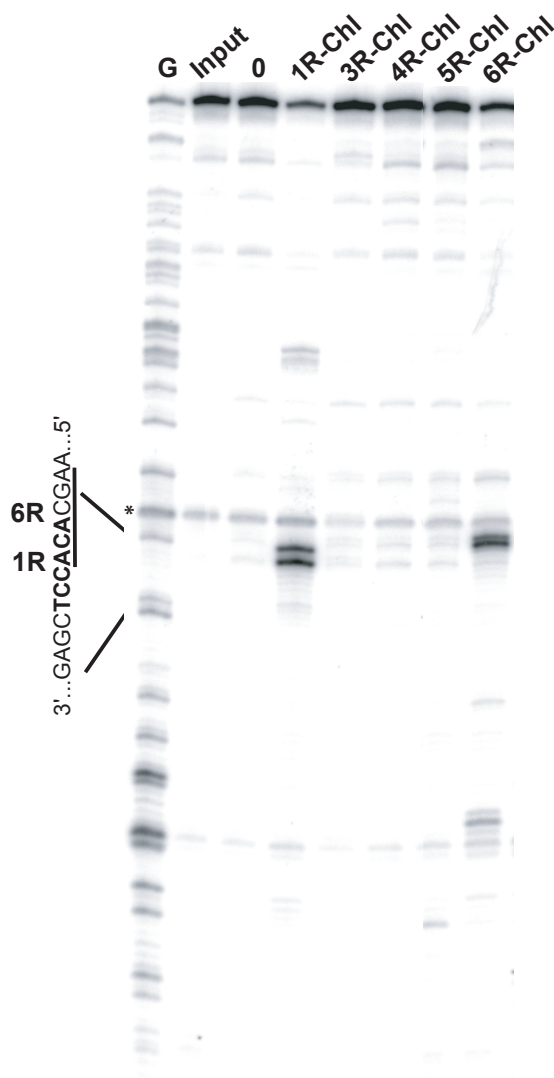


(continued from previous page)

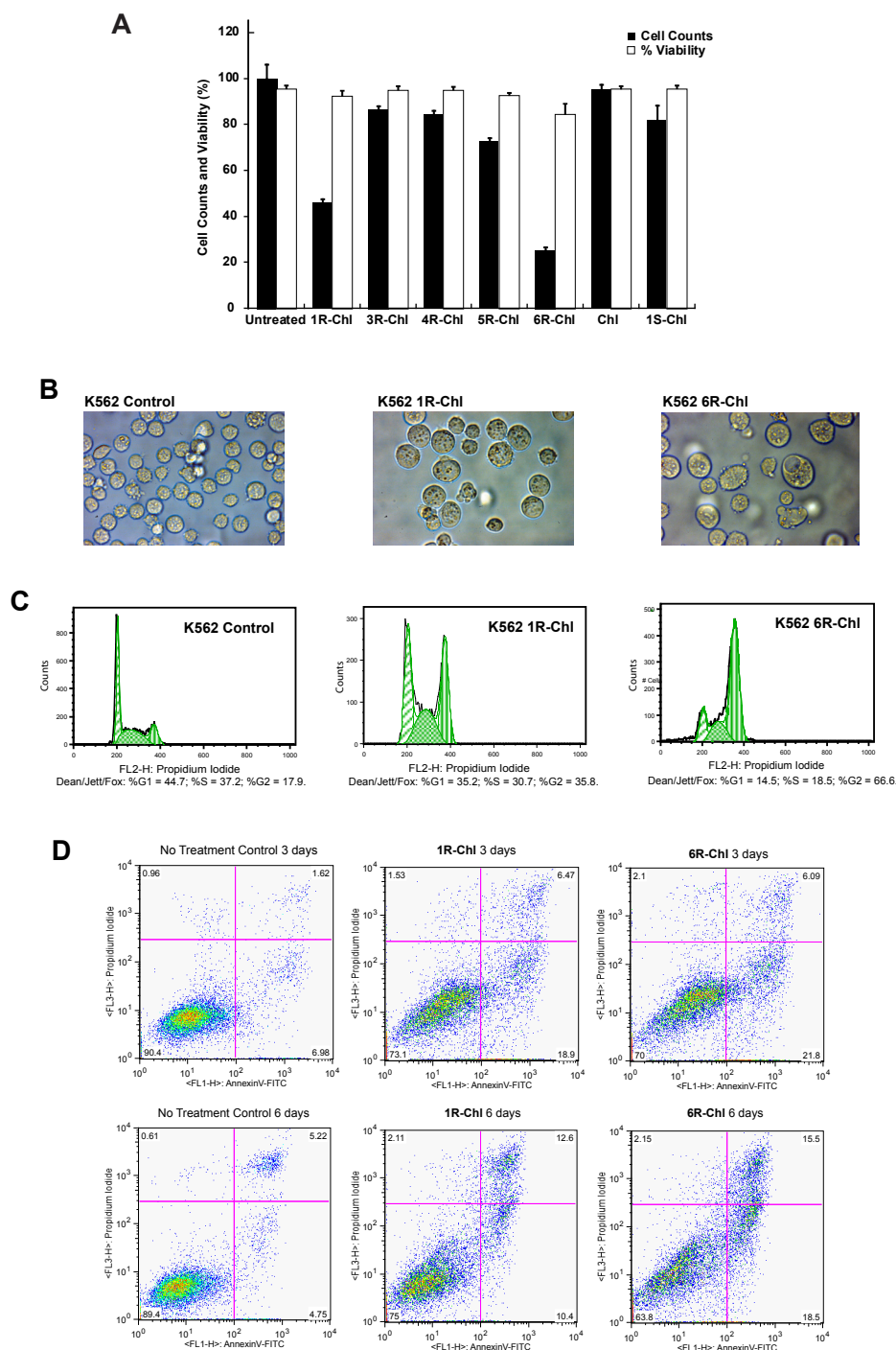
titration experiments; standard deviations are shown in parentheses (bottom). (C) Thermal cleavage assay experiments on pMFST6 with **4R-Chl** and **6R-Chl**. Experiments were performed at 1  $\mu$ M, 100 nM, and 10 nM for each polyamide. Putative major sites of alkylation on the resolved portion of pMFST6 are indicated by arrows under the sequences for **4R-Chl** (left) and **6R-Chl** (right). Lanes irrelevant to the figure have been cropped between lanes 5 and 6.

The ability of each polyamide-Chl conjugate to alkylate nucleotides adjacent to its match binding sites on synthetic DNA constructs was also evaluated. Alkylation profiles of polyamides **1R-Chl**, **3R-Chl**, **5R-Chl**, and **1S-Chl** have been previously reported;<sup>33-35</sup> polyamides **4R-Chl** and **6R-Chl** were analyzed on plasmid insert pMFST6 (Figure IIC.2C). All polyamides alkylated DNA in the nanomolar concentration range. H4c gene-specific DNA alkylation activities were monitored with the polyamide-Chl conjugates on a PCR product derived from the H4c gene. Significantly, only **1R-Chl** and **6R-Chl** have binding sites within the H4c gene and only these compounds effectively alkylate the H4c PCR product *in vitro* (Figure IIC.3). **6R-Chl** targets the DNA sequence 5'-WGWGCW-3'; the sequence 5'-TGTGCT-3' is found both adjacent to and overlapping the binding site for **1R-Chl** on the top strand of the H4c and H4k/j genes (5'-AGGTGT-3', Figure IIC.1B).

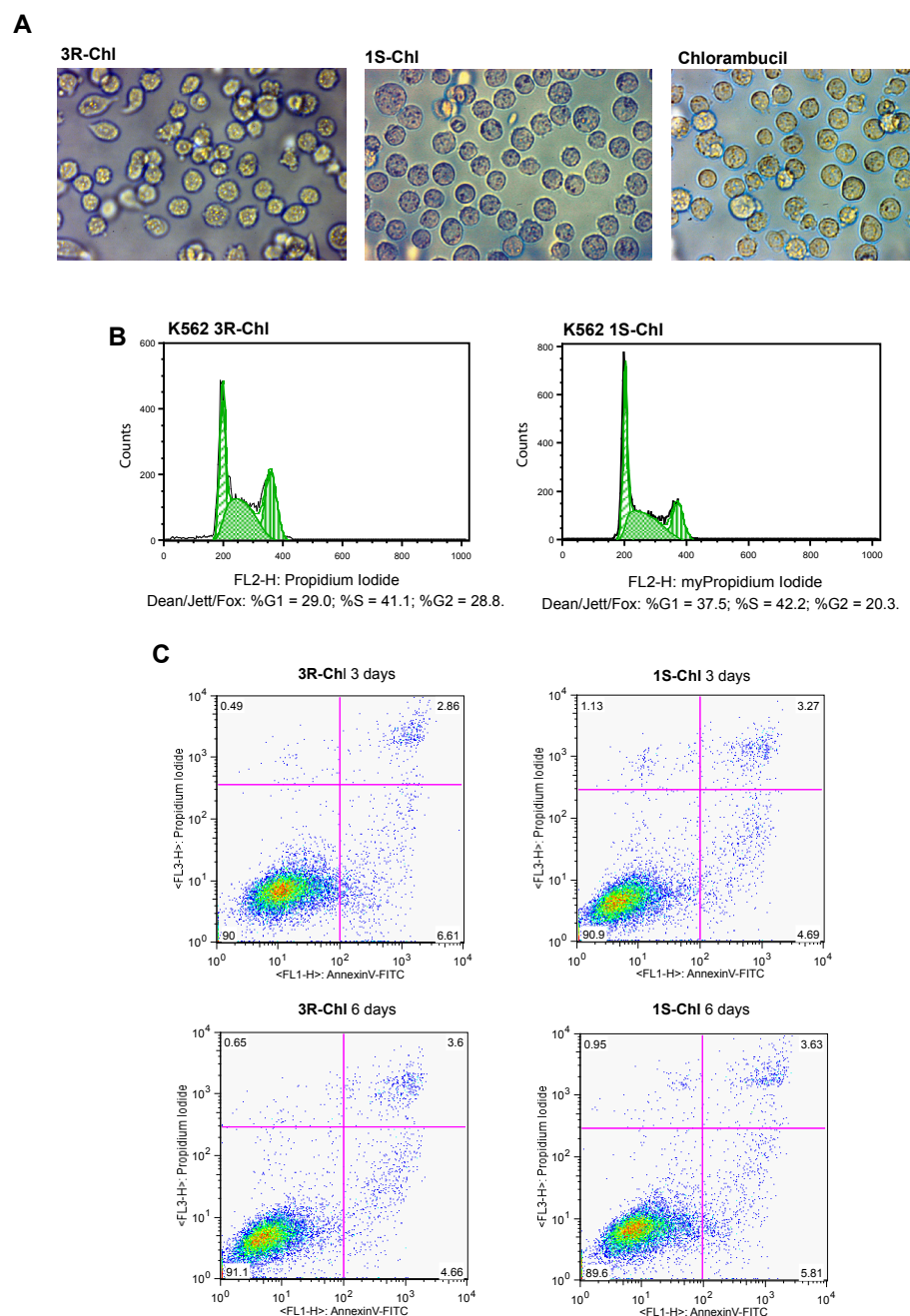
Polyamide-Chl DNA alkylators were then tested for effects on growth and morphology of K562 cells. Only cells treated with either **1R-Chl** or **6R-Chl** exhibited growth inhibition; all others were inactive against K562 proliferation (Table IIC.1, Figure IIC.4A). These molecules also caused an enlargement of cell volume with an  $IC_{50}$  of ~250 nM and induced G2/M cell cycle arrest, effects not observed with other molecules (Figure IIC.4B, 4C, 5A, and 5B). Following treatment with **1R-Chl** or **6R-Chl** for 3 to 6 days, an increase in the number of apoptotic cells was observed via annexin V-FITC/propidium iodine FACS analysis (Figure IIC.4C). Other polyamide-Chl conjugates had no effect on apoptosis (Figure IIC.5C).



**Figure IIC.3.** Alkylation of the H4c gene *in vitro*. Approximately 10 ng of a 240 bp PCR product derived from the H4c gene was incubated for 16 h at 37 °C in 100  $\mu$ L of 20 mM NaCl, 10 mM Tris-Cl, pH 7.4, with the indicated polyamide-Chl conjugates at 100 nM concentration, followed by thermal cleavage and primer extension. The radiolabeled primer interrogates alkylation events on the bottom, coding strand of DNA. Input DNA was not treated with polyamide or heat. '0' indicates a no polyamide control, and 'G' denotes a guanine-only sequencing reaction. Binding sites for **1R-Chl** (bold) and **6R-Chl** (underlined) are shown at the left. **1R-Chl** alkylates the adenine at the 5' end of its binding site and an adjacent guanine two bases 5' of the polyamide site, on the bottom strand (3'-TCCACACG-5', where the alkylated bases are shown in bold). **6R-Chl** alkylates two adenines at the 5' end of the binding site on the bottom strand (3'-ACACGAA-5', where the alkylated A's are shown in bold). Non-specific background bands are present in all lanes, including the input DNA, which are unrelated to the polyamides.



**Figure IIC.4.** Effects of polyamides in cell culture. (A) Effect of polyamides on growth and viability of K562 cells. Cells were incubated with the indicated polyamides at 250 nM for 3 days. Only **1R-Chl** and **6R-Chl** affected cell proliferation. (B) Phase microscopy images illustrate that **1R-Chl** and **6R-Chl** induce an increase in cell volume at 250nM after 3 days. (C) G2/M cell cycle arrest is observed for **1R-Chl** and **6R-Chl** after 24 h of treatment with 250 nM polyamide (top). **6R-Chl** caused a greater G2/M cycle arrest than **1R-Chl**. (D) AnnexinV/PI staining of **1R-Chl** and **6R-Chl** treated cells after 3 and 6 days. Treatment of cells with 250 nM **1R-Chl** or **6R-Chl** resulted in increased AnnexinV-positive and PI-negative populations and AnnexinV-positive and PI-positive populations.



**Figure IIC.5.** Analysis of inactive polyamides in K562 cells. (A) Phase microscopy images illustrate that **3R-Chl**, **1S-Chl**, and chlorambucil did not induce increases in cell volume at 250nM after 3 days. (B) No significant cell cycle perturbation (i.e. G2/M cycle arrest) is observed for **3R-Chl** and **1S-Chl** after 24 h of treatment with 250 nM polyamide. (C) AnnexinV/PI staining of **3R-Chl** and **1S-Chl** treated cells after 3 and 6 days (at 250 nM); neither molecule resulted in any effect.

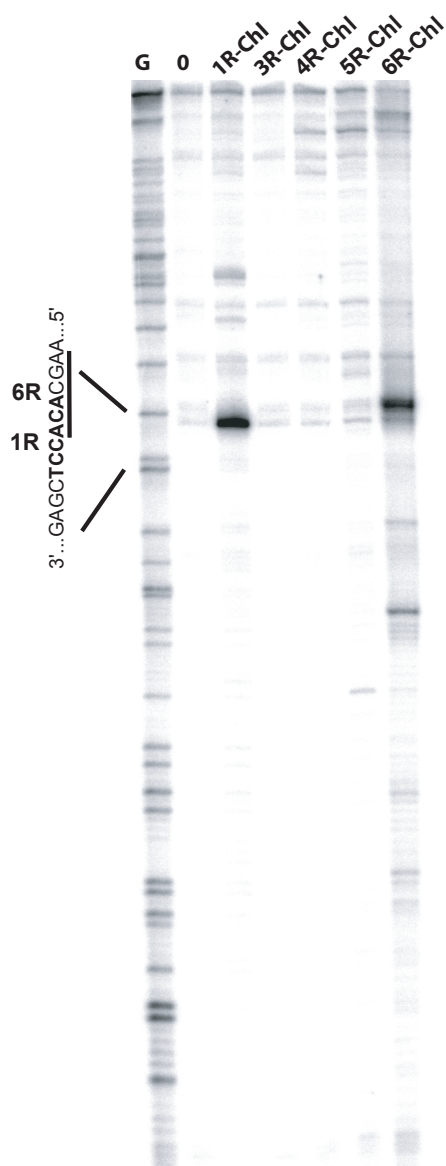
### Alkylation of the histone H4 genes by 1R-Chl and 6R-Chl in K562 cells

Ligation-mediated PCR (LM-PCR)<sup>36</sup> was used to determine sites of alkylation by **1R-Chl** and **6R-Chl** in the H4c gene in K562 cells. Cells were incubated with each of the hairpin polyamide-Chl conjugates for 24 h at a concentration shown to cause growth arrest with **1R-Chl** and **6R-Chl** (250 nM). After purification of genomic DNA and digestion with the restriction enzyme DraI, the DNA was heated to induce strand breakage at sites of alkylation.<sup>29</sup> LM-PCR with H4c gene-specific primers demonstrated that both **1R-Chl** and **6R-Chl** alkylate the H4c gene in cultured K562 cells. Consistent with *in vitro* alkylation results, other molecules in the library failed to alkylate the H4c gene *in vivo* (Figure IIC.6).

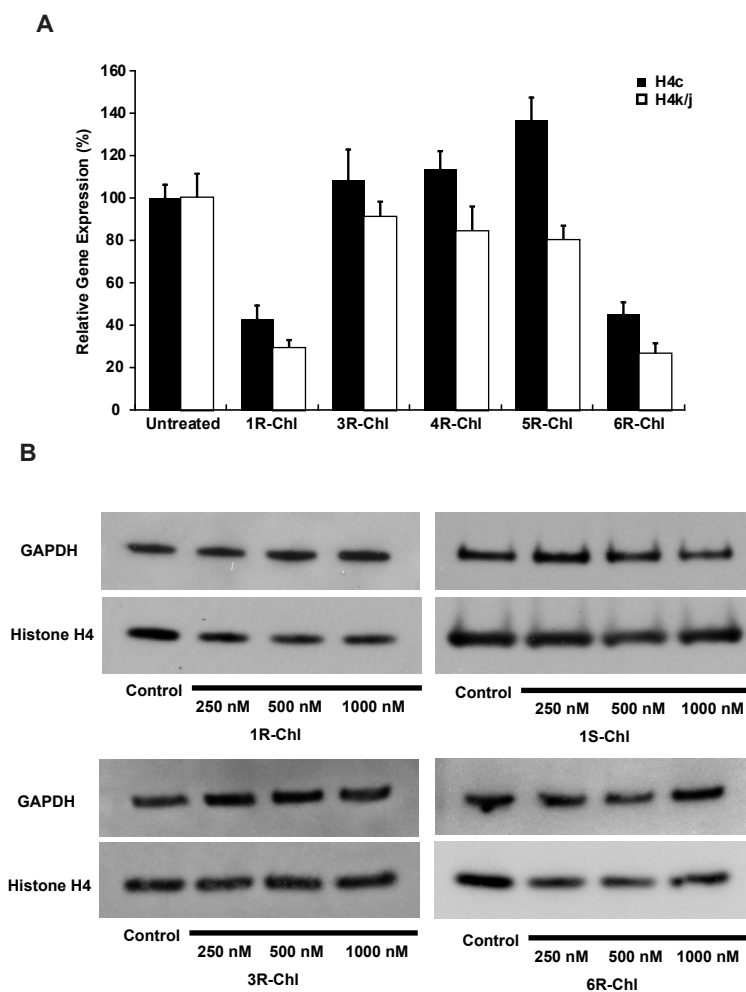
### Global gene expression, RT-PCR and Western blot analyses of 1R-Chl in K562 cells

Microarray studies have been previously conducted on SW620 cells treated with **1R-Chl**.<sup>32</sup> To determine whether **1R-Chl** treatment would cause similar changes to global gene expression in K562 cells, Affymetrix gene chip analysis was performed on K562 cells treated with **1R-Chl**. Through class comparison analysis, **1R-Chl** affected the levels of transcription of 156 genes: 60 genes were down-regulated and 96 genes were up-regulated ( $p = 0.00125$ ). Quantitative real-time PCR was used to verify the results of the expression arrays for several selected genes. The expression of histone H4 mRNAs H4c and H4k/j were decreased more than 60% compared to untreated control cells (Figure IIC.7A). Identical DNA binding and alkylation sites for **1R-Chl** are present in the H4c, H4k, and H4j genes. Western blot analysis for total histone H4 protein in whole cell lysate shows significant down-regulation of total H4 protein levels after 24 h of incubation with **1R-Chl** (Figure IIC.7B).

Through Affymetrix expression pathway analysis of **1R-Chl** treated K562 cells vs. untreated control ( $p = 0.001$ , MIT Board protocol), 2 significant pathways were identified: 87 genes in chemicalPathway\_h and 55 genes in mitochondriapathway\_h. Both pathways



**Figure IIC.6.** Alkylation of the H4c gene in K562 cells. Cells were incubated with the indicated polyamide-Chl conjugates for 24 h, followed by DNA purification and LM-PCR. The radiolabeled primer interrogates alkylation events on the bottom, coding strand of DNA. “0” indicates a no polyamide control, and “G” denotes a guanine-only sequencing reaction performed on a PCR product from the H4c gene, followed by LM-PCR. Binding sites for **1R-Chl** (bold) and 6R-Chl (underlined) are shown at left. Both molecules alkylate the adenine at the 5’ end of their respective binding site. Non-specific background bands appear in all lanes and are unrelated to the polyamides.



**Figure IIC.7.** Effect of polyamides on histone H4c and H4k/j transcript expression and Western blot analysis. (A) Chart showing polyamide effect on H4c and H4k/j expression; only **1R-Chl** and **6R-Chl** caused gene down-regulation. (B) K562 cells were treated with **1R-Chl** and **6R-Chl**, at the indicated concentrations for 24 h, and histone H4 and GAPDH protein levels were assayed by Western blot analysis (top left and bottom right, respectively). Analysis of treatment with **1S-Chl** and **3R-Chl** (top right and bottom left, respectively) did not show significant decreases in histone H4 levels.

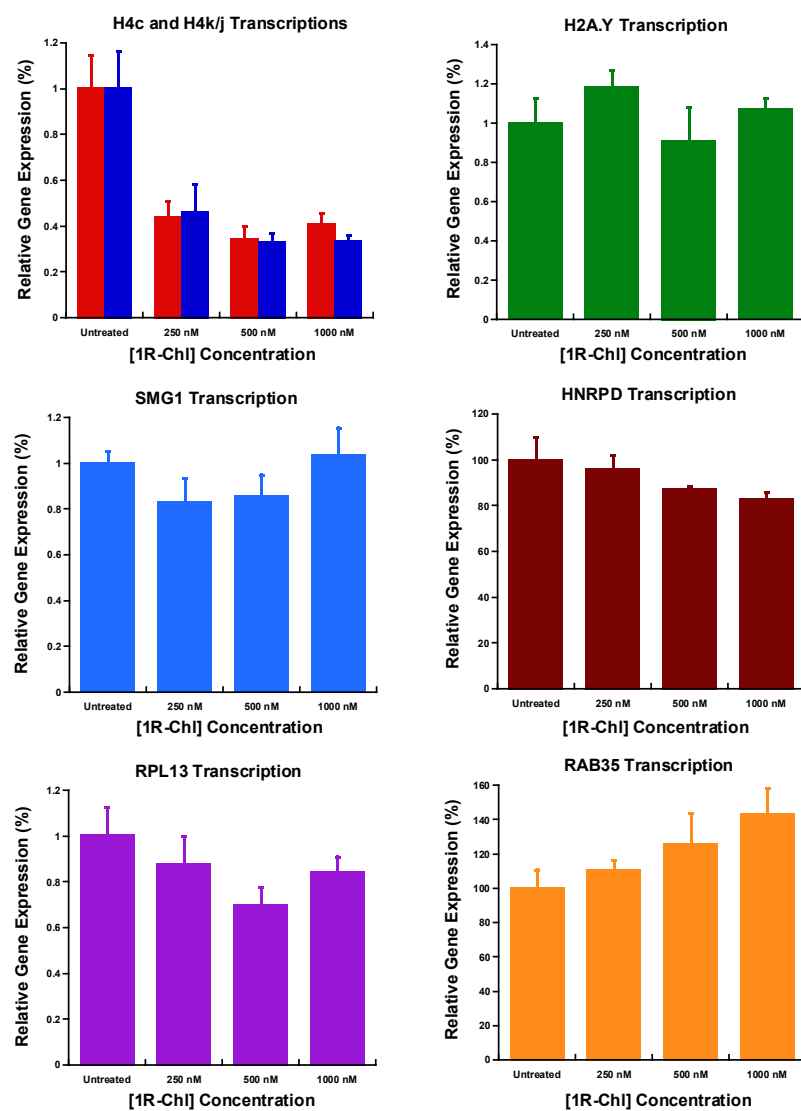
suggest initiation of apoptotic signaling in response to DNA damage with several known genes identified (i.e., BAX, BID, STAT1, ATM, BCL2, CASP3, CASP6, CASP7, CASP8, CASP9). These results correlate with up-regulated genes found in the expression array class comparison with increases in CHEK2 and BID expression levels. Surprisingly, several of the most down-regulated and highly significant genes, based on *p*-values and extent of down-regulation in microarray analysis (i.e. SMG1, RPL13, HNRPD, RAB35, and H2AFY) appeared not to be significantly affected by **1R-Chl** when analyzed by real-time qRT-PCR (Figure IIC.8).

### **1R-Chl and 6R-Chl arrest the growth of K562 xenografts**

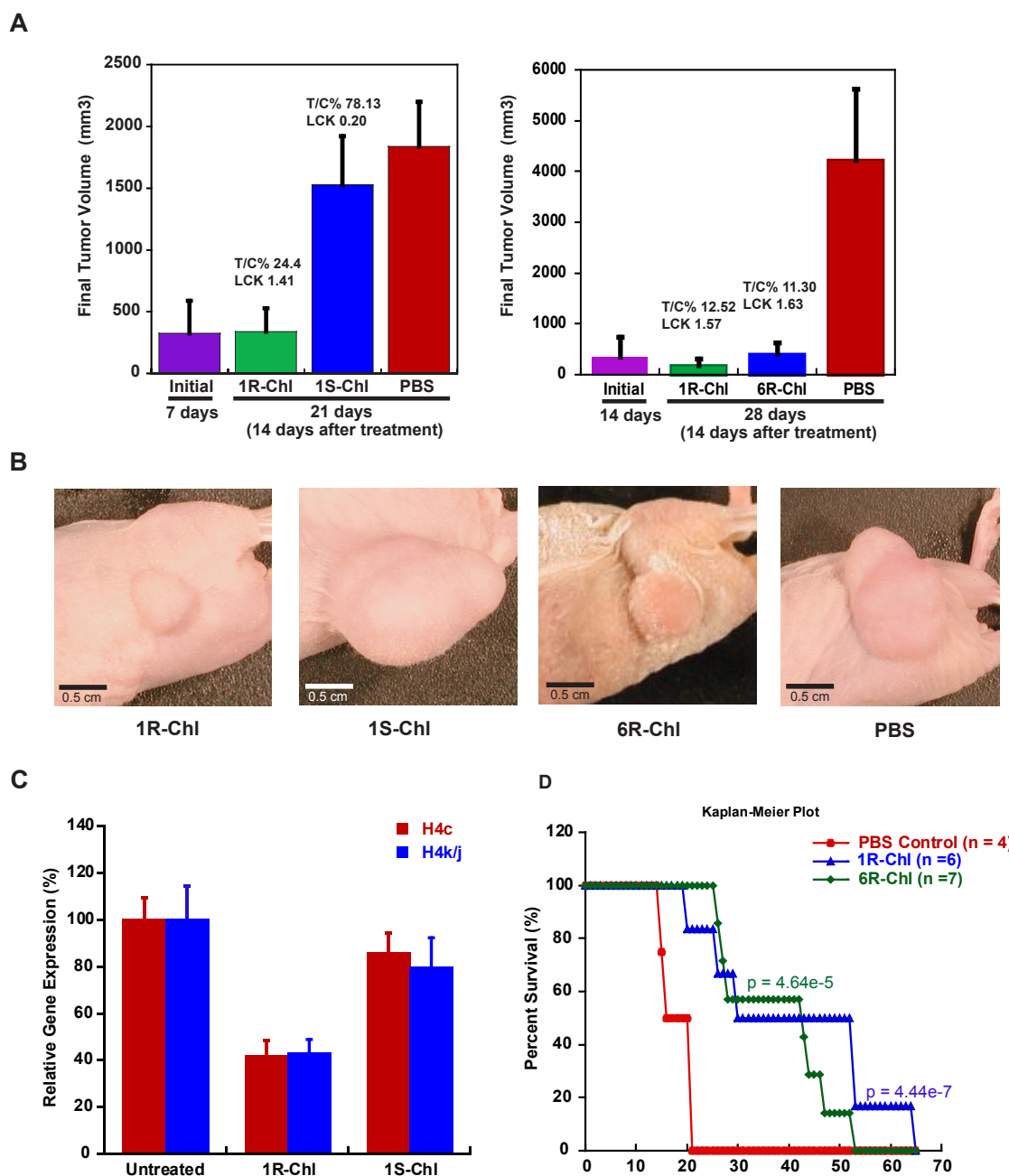
K562 cells were injected into athymic nude mice, and tumors were staged to >100 mm<sup>3</sup> (approximately 7 to 14 days post injection of cells). In the first experiment, mice were then intravenously injected with three doses of **1R-Chl**, **1S-Chl**, or PBS vehicle control (Figure IIC.9A, left); in a second experiment mice were injected with **1R-Chl**, **6R-Chl**, or PBS vehicle (Figure IIC.9A, right). The molecule **1S-Chl** was used as a control polyamide; although it targets the same sequence as **1R-Chl**, the opposite stereochemistry of its turn unit results in a greatly reduced ability to bind and alkylate DNA.<sup>32,33</sup> The dosage routine was 7.5 mg/kg per injection and the polyamides were administered on treatment days 0, 2, and 5. The treatment dose of **1R-Chl** was determined based on the LD<sub>20</sub> of Tallimustine and Brostallicin, both of which have shown severe dose limiting toxicities in the murine model and during phase II clinical trials.<sup>9,37</sup>

Mice treated with **1R-Chl** and **6R-Chl** showed immediate growth regression of the tumor xenografts, while **1S-Chl** and PBS vehicle had no effect on tumor growth (Figure IIC.9A and 9B). Importantly, all mice treated with either **1R-Chl** or **6R-Chl** appeared healthy with minimal weight loss, while all **1S-Chl** and PBS vehicle treated mice showed signs of wasting with significant weight loss due to exponential tumor growth. No obvious toxicity was associated with polyamide treatment. Based on the T/C% (tumor volume





**Figure IIC.8.** Real-time qRT-PCR of several genes identified in Affymetrix microarray experiments: H4c, H4k/j, SMG1, H2A.Y, HNRPD, and RAB35. All transcripts were interrogated with appropriate primers against GAPDH as a normalization control. Only H4c (Red) and H4k/j (Blue) are down-regulated significantly from 250 nM to 1000 nM within 24 h while the rest of the gene transcripts are without significance.



**Figure IIC.9.** Murine K562 xenograft studies. (A) Athymic nude mice were injected with K562 cells and tumors were allowed to develop for either 7 days (left) or 14 days (right), after which mice were treated with the indicated polyamides; tumor volumes were determined 14 days later. (B) Representative photographs of **1R-Chl**, **6R-Chl**, **1S-Chl**, and PBS-treated animals. (C) mRNA was isolated from tumor xenografts after treatment with **1R-Chl** and **1S-Chl**, 24 h after the last injection. Real-time qRT-PCR showed down-regulation of H4c and H4k/j transcripts *in vivo* by **1R-Chl**; **1S-Chl** did not significantly down-regulate either H4c or H4k/j. (D) Kaplan-Meier survival plot showing extended lifespan of **1R-Chl** and **6R-Chl** treated mice. The life extensions by both **1R-Chl** and **6R-Chl** are approximately 2 weeks with  $p = 4.44 \times 10^{-7}$  and  $4.64 \times 10^{-5}$ , respectively.

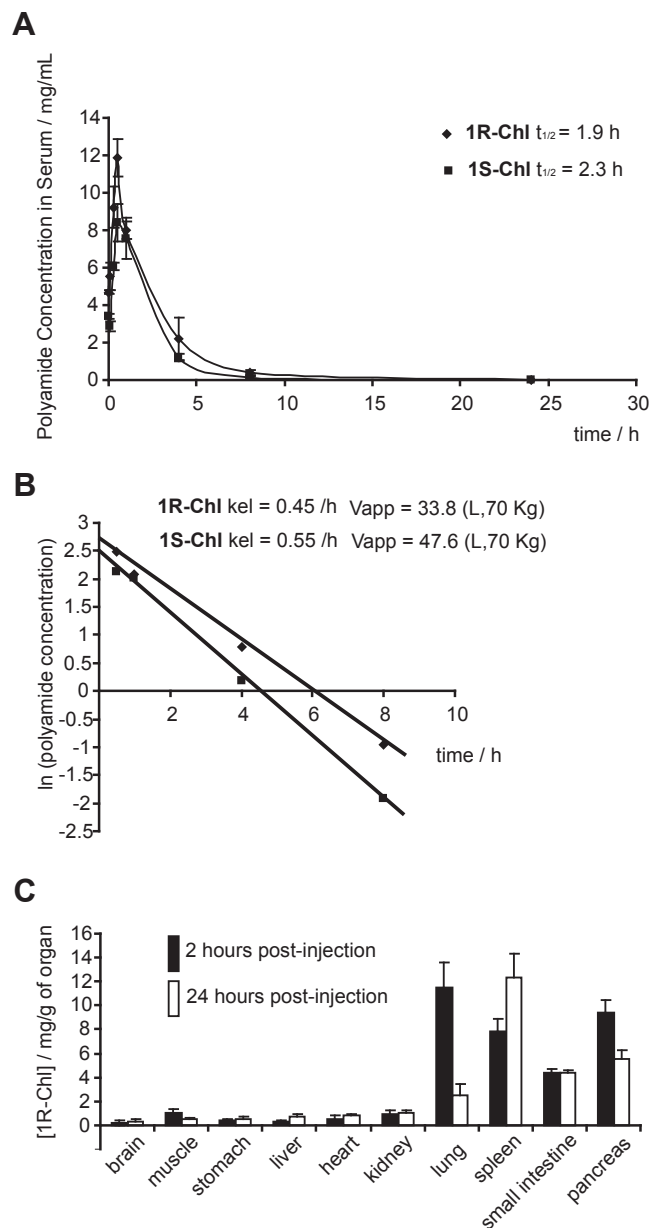
analysis) and LCK values (tumor growth delay; Log Cell Kill), **1R-Chl** and **6R-Chl** treatments are highly effective: according to the National Cancer Institute standards,  $T/C < 42\%$  and  $LCK > 0.7$  are minimal levels for activity.<sup>37</sup> For the two experiments conducted, the  $T/C\%$  for **1R-Chl** treated mice versus untreated control were 24.4% and 12.52%, and LCK was 1.41 and 1.57, respectively. For **6R-Chl** the  $T/C\%$  was 11.30% and LCK was 1.63. In contrast, treatment with **1S-Chl** gives results outside of the therapeutic range ( $T/C\% = 78.13$  and  $LCK = 0.20$ ).

24 h following the final **1R-Chl** treatment, K562 xenograft tumors were dissected and their mRNA was isolated for qRT-PCR analysis. Both H4c and H4k/j transcripts were observed to be down-regulated in tumors from **1R-Chl** treated mice. Analysis of tumor samples from **1S-Chl** and PBS treated mice revealed no effect on histone H4c/k/j mRNA levels (Figure IIC.9C), consistent with *in vitro* results.

A Kaplan-Meier survival plot for mice treated with **1R-Chl** or **6R-Chl** versus PBS shows a significant extension in lifespan, even after only one dosing regimen (Figure IIC.9D). We expect that prolonged administration of these polyamide-Chl conjugates would extend lifespan even further, and will be determined in future studies.

### Pharmacokinetic parameters and pharmacotoxicity of **1R-Chl**.

To predict the behavior and likely dosage regime of the polyamides in humans, pharmacokinetic parameters for **1R-Chl** and **1S-Chl** were determined in Balb/c mice. Mass spectrometry was used to determine polyamide concentration in serum following injection of **1R-Chl** or **1S-Chl**. Both compounds appeared to exhibit first-order decay kinetics after a brief lag time (Figure IIC.10A), indicating that the rate of elimination from serum is proportional to the concentration of the compound in serum at a particular time-point. From these results, we determined that both **1R-Chl** and **1S-Chl** have  $t_{1/2}$  values of  $\sim 2$  h. The constant of elimination ( $k_{el}$ ) was determined by plotting the natural logarithm of **1R-Chl** serum concentration vs. time (during the clearance phase), and calculating the



**Figure IIC.10.** Pharmacokinetic parameters and biodistribution of **1R-Chl**. (A) Concentration (in mg/mL, y-axis) of **1R-Chl** (rhombus) and **1S-Chl** (squares) in serum over time. Mice were injected with either **1R-Chl** or **1S-Chl** (100 nmol, equivalent to 7.5 mg/kg) and blood extracted at indicated time-points. Amounts of polyamide were determined by mass spectrometry, and half-lives ( $t_{1/2}$ ) for **1R-Chl** and **1S-Chl** calculated. (B) Graph depicting the natural logarithm (ln) of **1R-Chl** (rhombus) and **1S-Chl** (squares) concentration versus time. The y-intercept represents the theoretical concentration at  $t = 0$ , where the constant of elimination and the apparent volume of distribution were calculated. (C) Graph depicting the biodistribution of **1R-Chl**. Mice were treated with **1R-Chl** (500 nmol = 37.5 mg/kg), and euthanized either 2 h (black bars) or 24 h (white bars) post-injection. Major organs were harvested, and a weighed portion of each homogenized and sonicated. **1R-Chl** concentration was determined by mass spectrometry.

slope of the best fit straight line;  $k_{el}$  was found to be  $\sim 0.45\text{ h}^{-1}$ . The concentration of the compound achieved in blood at  $t = 0$  was determined by extrapolating the line to the y-intercept. Volume of distribution ( $V_d$ ) was calculated by dividing the dose injected by this theoretical initial concentration ( $t_0$ ). **1R-Chl** was predicted to have a  $V_d$  of  $\sim 0.48\text{ L/kg}$ , corresponding to  $\sim 33.8\text{ L}$  for a 70 kg human (Figure IIC.10B). PK parameters for **1S-Chl** were also determined and are similar to those of **1R-Chl**:  $k_{el} \sim 0.55\text{ h}^{-1}$ ;  $V_d \sim 0.68\text{ L/kg}$ ;  $V_d \sim 47.62$  for a 70 kg human.

The distribution of **1R-Chl** in various mouse organs was determined by mass spectrometry (Figure IIC.10C). Following a 500 nmol dose, equal amounts of **1R-Chl** were detected in the brain, muscle, stomach, liver, heart and kidney; it is likely that **1R-Chl** crosses the blood-brain barrier. Extensive amounts of **1R-Chl** were found in the lung, spleen, small intestine, and pancreas after 2 h post-injection and at  $t_{1/2}$ , as well as 24 h post-injection, at which point **1R-Chl** is undetectable in serum (Figure IIC.10A). These results indicate that **1R-Chl** can be readily eliminated from serum, and persist in peripheral organs after tissue entry.

### **Histone genes as primary targets of polyamide-chl conjugates leading to growth inhibition of cancer cells.**

In the library of polyamide-Chl conjugates described herein, **1R-Chl** and **6R-Chl** are the only molecules expected to bind and alkylate within histone H4c/k/j genes. Sites of alkylation within the H4c gene are on the template strand, at the 5' ends of the **1R-Chl** and **6R-Chl** targeted sequences (Figure IIC.1, 3, and 6). Alkylation of this strand would be predicted to block transcription elongation by RNA polymerase II. Reporter gene assays have shown that alkylation of the template strand produces a more pronounced effect on transcription inhibition than on the non-coding strand.<sup>38</sup> Real-time qRT-PCR and Western blot analyses confirm that K562 cells treated with **1R-Chl** or **6R-Chl** (250 nM for 24 h) exhibit down-regulation of H4c and H4k/j transcripts and total histone H4 protein levels.

Other polyamide-Chl conjugates do not alkylate H4 genes *in vitro* or in living cells and neither affect H4c and H4k/j mRNA and histone H4 levels, nor inhibit K562 cell growth.

Previous studies involving **1R-Chl** have shown that down-regulation of the H4c gene in SW620 cells is required for growth inhibition, but changes in H4k/j transcription were not observed.<sup>32</sup> Further inspection of Affymetrix microarray data for SW620 cells revealed that H4k/j genes are not highly expressed in SW620 cells relative to H4c. However, K562 cells have similar H4c and H4k/j expression levels as confirmed by real-time qRT-PCR. The primary targets of down-regulation by **1R-Chl** in K562 cells were the histone H4 genes, H4c and H4j/k. The only gene down-regulated by **1R-Chl** in both K562 and SW620 cells is H4c, indicating it to be the key target that elicits growth effects. Furthermore, only cultured cell lines that express the histone H4c gene at high levels have been shown to be affected by **1R-Chl**.<sup>32,35</sup>

### **Polyamide-Chl conjugates block tumor progression with little animal toxicity**

Polyamides **1R-Chl** and **6R-Chl** are the first sequence-specific DNA alkylators reported to show little animal toxicity at mid to high mg/kg range with three consecutive injections every other day over 5 days. This dosing regimen is highly effective in blocking the growth of established tumors (Figure IIC.9). No behavior or significant weight changes were observed during experiments with normal Balb/c or xenograft mice. In comparison, Tallimustine and Brostallicin, both of which are micromolar DNA alkylators targeting AT-rich sequences, have shown significant animal toxicity at 4 mg/kg and 0.8 mg/kg respectively.<sup>9,37</sup>

We have shown here that **1R-Chl** possesses desirable PK parameters: it has a  $t_{1/2}$  of ~ 2 hours – enough time to allow the drug to effectively reach its target while minimizing prolonged presence in the blood stream, which can result in toxic effects. The constant of elimination of **1R-Chl** is relatively low (~0.45/h) compared to those of most anti-cancer drugs, indicating that should a steady concentration of the drug be necessary to achieve

efficacy, a regime of constant infusion may be considered. The  $k_{el}$  of **1R-Chl** is similar to that of Capecitabine (0.5/h), a non-alkylating anti-cancer drug used to treat breast cancer. Also,  $V_d > 30$  for **1R-Chl**, indicating that it is very well distributed in the body, an observation which is consistent with the presence of the molecule in the brain, muscle, stomach, liver, heart, and kidney in our biodistribution analysis (Figure IIC.10C). The PK parameters for **1R-Chl** are close to those of some commonly used chemotherapeutic drugs, including both alkylating agents and other types of compounds.

## Conclusion

The development of the selective p210 Bcr-Abl tyrosine kinase inhibitor Gleevec (STI 571, Imatinib mesylate, Novartis) as a directed therapeutic for CML has been a major advance in cancer therapy.<sup>39,40</sup> However, many Gleevec patients acquire resistance mutations to the drug or acquisition of Bcr-Abl-independent genetic abnormalities during the course of treatment.<sup>41-44</sup> Thus, development of additional therapeutic approaches for CML and other malignancies is worthwhile. Here we describe two sequence-specific DNA alkylators, **1R-Chl** and **6R-Chl**, that are capable of inhibiting K562 CML cell growth both *in vitro* and *in vivo*. **1R-Chl** and **6R-Chl** treatments are also highly effective in blocking K562 xenograft growth, with high dose tolerance in the murine model. Based on these observations and the finding that **1R-Chl** blocks SW620 xenograft growth in nude mice,<sup>32</sup> and the growth of two other xenograft models (Calu-1 lung cancer cells and 22Rv1 prostate cancer cells, data not shown), polyamide-Chl conjugates appear to be promising cancer therapeutics. The inhibitory mechanism of **1R-Chl** and **6R-Chl** in K562 cells is likely to be independent of Bcr-Abl because no significant down-regulation of Bcr-Abl transcripts was observed in either Affymatrix genechip analysis or real time qRT-PCR (data not shown). These results suggest that **1R-Chl** maybe an effective treatment alternative for CML or may be useful in combination with Gleevec.

## Materials and Methods

### Synthesis and characterization of pyrrole-imidazole polyamides.

Pyrrole-imidazole (Py-Im) polyamides were synthesized by standard solid phase methods,<sup>45</sup> using  $\alpha$ -(*R*)- or  $\alpha$ -(*S*)-2,4-diaminobutyric acid as the hairpin turn unit.<sup>33</sup> Polyamide-Chl conjugates were generated as previously described,<sup>29</sup> whereby the carboxylic acid of Chl (Sigma-Aldrich) is activated and coupled to the free amine of the hairpin turn. The identity and purity of the compounds were established by mass spectrometry analysis (MALDI-TOF-MS and ES-MS) and analytical HPLC, respectively.

Binding affinities of the parent polyamides (which lack chlorambucil) for target match and mismatch sites were determined by quantitative DNase I footprinting,<sup>46</sup> using a radiolabeled DNA fragment of the plasmid pMFST6. The plasmid was constructed by annealing the oligonucleotide pair: 5' - A G C T G T A G T A T C T A T A - G T G C T T A C T A T C T A T A G T C C T T A C T A T C T A T A G C T C T T A C T A T C T A T A G C T G T T A - C T A T C T A T A C T - A T C T A C - 3' and 5' - C A T C A T A G A T A T C A C G A A T G A T A G A T A T C A - G G A A T G A T A G A T A T C G A G A A T A T A G A T A T C G A C A A T G A T A G A T A G A T A G A T G - C T A G - 3'. Annealed nucleotides were ligated into the *Bam*HI/*Hind*III restriction fragment of pUC19 using T4 DNA ligase, and the plasmid was transformed into *Escherichia coli* JM109 competent cells. Ampicillin-resistant white colonies were selected from 25 mL Luria-Bertani agar plates containing 50 mg/mL ampicillin treated with XGAL and isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) solutions and grown overnight at 37 °C. Cells were harvested the following day, and purification of the plasmid was performed with a Wizard Plus Midiprep DNA purification kit (Promega). Alkylation experiments performed on the pMFST6 insert were performed as previously described.<sup>29</sup> Alkylation experiments were also performed on a 240 bp region of the H4c mRNA-coding sequence, which was amplified from genomic DNA with the following PCR primers: 5' - G T G C T A A G C G C C A T C G T A A G - 3' and 5' - C C C T G A C G T T T T A G G G C A T A - 3'. These experiments were conducted using 10 ng of the PCR product incubated for 16 h at 37 °C in 100  $\mu$ L of 20 mM NaCl, 10 mM Tris-Cl,



pH 7.4, with each of the polyamide-Chl conjugates at 10, 100 and 1000 nM concentration, followed by thermal cleavage and primer extension labeling, as described.<sup>32</sup>

### **Cell lines and cell viability assays**

The human CML lymphoblast cell line K562 (purchased from ATCC), which contains the b3a2 Bcr-Abl translocation, was used in this study. Cells were grown and maintained in RPMI 1640 medium containing 10% FBS, under standard mammalian cell culture conditions as recommended by the ATCC. Direct phase contrast microscopic visualization was used to monitor the effects of polyamide-Chl conjugates on cell growth rates and cell morphology. Promega CellTiter 96 Aqueous One Solution Cell Proliferation assay (Promega, WI; utilizing [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H tetrazolium (MTS) conversion to formazan to examine mitochondrial activity), Trypan blue exclusion, and Annexin V-FITC/propidium iodide (PI) apoptosis staining (BD Pharmingen, CA) were used to determine cell proliferation, viability, and initiation of apoptosis, respectively.

### **Cell cycle analysis**

The effects of polyamide-Chl conjugates on cell cycle progression were monitored by flow cytometry analysis in the Scripps FACS core facility. Polyamide-treated cells (250 nM of polyamide in culture media for 24 h) were collected by centrifugation (200X g for 5 min). Cell pellets were re-suspended in 500  $\mu$ L of PBS and fixed with addition of 4.5 mL of pre-chilled 70% ethanol, stained with propidium iodide (50  $\mu$ g/mL), and analyzed for DNA content, reflecting the fraction of cells at each point in the cell cycle ( $G_0$ /G1, S, and G2/M). Cells with less than a 2C DNA content are indicative of DNA fragmentation and apoptosis.

### **Ligation-mediated PCR**

K562 CML cells were incubated in culture medium for 24 h with each of the polyamide-Chl conjugates at a 250 nM concentration, followed by digestion of purified DNA with DraI, thermal cleavage and ligation-mediated PCR, using the primers for the H4c gene listed above, as previously described.<sup>32</sup>

### **Affymetrix expression array experiments and data analysis**

K562 cells were incubated with **1R-Chl** (at 250 nM) or in the absence of polyamide, in triplicate for 24 h before RNA purification and microarray analysis at The Scripps Research Institute microarray facility. Affymetrix U133A Plus 2.0 GeneChips were hybridized in groups of three for each of the two groups. The Affymetrix probe set data were imported into BRB Arraytools (3.5.0 Beta 2), selecting the U133 chips used in the experiment and leaving all filters off. For class comparison analysis, **1R-Chl** treated vs. untreated controls, all three replicates of each group were used, and the random variance model was selected, with the univariate significant threshold set to 0.00125. For the pathway analysis, MIT Board pathway library was used. Random variance model was used with LS permutation test or KS permutation test significance set at the nominal 0.001 level. Microarray data has been deposited at Gene Expression Omnibus (accession number GSE8832).

### **Real-time quantitative RT-PCR**

RNA from polyamide-treated cells was extracted using the Absolutely RNA Miniprep kit (Stratagen, CA). RT-PCR was performed using iScript One-Step TR-PCR kit with SYBR green (Bio-Rad Laboratories, CA) in accordance with the manufacturer's instructions. Levels of H4c, H4k/j, SMG1, RPL13, HNRPD, RAB35, and H2AFY transcripts were quantified by amplifying a segment of their respective mRNAs with appropriate primer sets (Table IIC.2). The reverse-transcription reaction was carried out at 50 °C for 10 min, followed by iTaq hot-start DNA polymerase activation by heating

at 95 °C for 15 min. Three-step cycling was performed: denaturation – 15 s at 95 °C, annealing – 30 s at 55 °C, and extension – 30 s at 72 °C, for 45 cycles. All gene expression levels were normalized by parallel amplification and quantification of mRNA levels for the housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH) mRNA, as an endogenous reference with the following primers: 5'-GAGTCAACGGATTTGGTCGT-3' and 5'-GAGGTCAATGAGGGGTCAT-3'.

---

#### **H4c Primers**

Forward 5'- GGG ATA ACA TCC AGG GCA TT- 3'

Reverse 5'- CCC TGA CGT TTT AGG GCA TA- 3'

#### **H4kj Primers**

Forward 5'- CGC CGT GAC CTA TAC AGA GC-3'

Reverse 5'- AAC CAC CGA AAC CGT AGA GG-3'

#### **SMG1 Primers**

Forward 5'- CAA GCG ATG TCA GCA GAT GT- 3'

Reverse 5'- TGC TGA CAA AAG CCA TTC AG- 3'

#### **RPL 13**

Forward 5'- GCA TTT CTG TGG ATC CGA GGA G-

Reverse 5'- CGA CTG ATT CCA AGT CCC C- 3'

#### **H2A.Y**

Forward 5'- CTG ACA GCG GAG ATT CTG G- 3'

Reverse 5'- CTT CAT CAT TGG CCA CAG C- 3'

#### **HNRPD**

Forward 5'- TTT TGT TGG TGG CCT TTC TC- 3'

Reverse 5'- GTT GTC CAT GGGGAG CTC TA- 3'

#### **RAB35**

Forward 5'- GCA TCC AGT TGT TCG AGA CC- 3'

Reverse 5'- GCC AGG TTG TCT TTC TTT GC- 3'

---

**Table IIC.2.** Primers used for real-time qRT-PCR. The primers are designed with primer3 and amplification efficiency is verified by titration of K562 mRNA in comparison to GAPDH primers.

**SDS-PAGE and Western blot analysis**

Equal numbers of cells were lysed with RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% deoxycholate, 1 mM EDTA) with 1 X complete mini protease inhibitor cocktail (Roche Diagnostics, IN) for 30 min at 4 °C followed by 15 s sonication pulses (Branson Sonifier-150 at 3 watts). Cell lysates were then centrifuged at 14,000 rpm (12,000 g) for 15 min and the supernatants removed and combined with LDS sample loading buffer (Invitrogen, CA). SDS-PAGE was performed with the Invitrogen NuPAGE system using 4-12% Bis-Tris gels and MES running buffer (Invitrogen, CA). Electrophoresis was carried out at 200 V for 45 min, and gel contents were transferred at 30 V for 1 h to a 0.2  $\mu$ m nitrocellulose membrane. Membranes were then blocked with 5% bovine serum albumin for 1 h at 4 °C and probed with histone H4 or GAPDH (Abcam, CA) primary antibodies. Protein-antibody complexes were then visualized by enhanced chemiluminescence using Amersham ECL system (GE Healthcare, UK), with either anti-rabbit or anti-goat horseradish peroxidase (HRP) conjugated secondary antibodies (Santa Cruz Biotechnology, CA).

**Xenograft studies of 1R-Chl, 6R-Chl, and 1S-Chl**

Female athymic nude mice were purchased from The Scripps Research Institute Division of Animal Resources. Experimental protocols were approved by the Scripps Institutional Animal Welfare Committee. K562 cells were suspended to 50 million cells/mL in Matrigel (BD Biosciences, CA), 0.2 mL of which was subcutaneously injected into the rear left flank of each mouse (6-8 wks of age). Mice were monitored and tumor sizes measured daily; tumor volumes were calculated as  $\frac{1}{2}$  length x (width)<sup>2</sup>. Tumors were staged for 7 to 14 days to enter growth phase. At this point compounds were administered at a dose of 7.5 mg/kg per injection, via the tail vein three times over a five-day period.

### Pharmacokinetic properties of **1R-Chl** and **1S-Chl**

Pharmacokinetic (PK) studies were performed in normal Balb/c mice in order to determine plasma levels of the compounds **1R-Chl** and **1S-Chl** over time, and to calculate their constant of elimination ( $k_{el}$ ), half-life ( $t_{1/2}$ ), and volume of distribution ( $V_d$ ). 48 female Balb/c mice were divided into three groups of 16 mice each. Group 1 received 100  $\mu$ L (100 nmol) of 1mM **1R-Chl**, group 2 received the same amount of **1S-Chl**, and group 3 received an equal volume of PBS (control) via I.V. bolus injection. 0.5 mL of blood was collected from each animal at specific time-points (0, 5, 15, 30 min and 1, 4, 8 and 24 h) in heparinized tubes, then allowed to coagulate at ambient temperature for 24 h, and centrifuged at 3,100 rpm for 10 min to separate the serum from the clot.

Concentrations of **1R-Chl** and **1S-Chl** in the serum were determined by mass spectrometry (see below). Each time-point had two animals per group. Data collected was plotted as concentration (y-axis) vs. time (x-axis), and subjected to first order kinetic analysis in order to calculate the PK parameters. From this graph,  $t_{1/2}$  was calculated, and from the natural logarithm ( $\ln$ ) of the concentration vs. time,  $k_{el}$  (slope) and  $V_d$  (dose/anti- $\ln$  y-intercept) were determined (y-intercept represents the theoretical concentration at  $t = 0$ ).  $V_d$  for humans was predicted by multiplying the  $V_d$  by 70 kg (average human weight).

**1R-Chl** distribution in the mouse body was analyzed as follows: mice were injected with 500 nmol of **1R-Chl** or PBS and euthanized after either 2 h or 24 h; organs were fixed in formalin. A small piece from each major organ was weighed, homogenized in PBS buffer, and sonicated by two pulses of 15 s each. **1R-Chl** concentration was determined by mass spectrometry (see below).

### Mass spectrometry determination of **1R-Chl** and **1S-Chl**

A control polyamide-Chl conjugate, with the sequence ImImPyIm- ( $R$ )<sup>Chl</sup> $\gamma$ -PyPyPyPy- $\beta$ -Dp, was used as an internal standard for concentration determination by

mass spectrometry. The molecular mass of this molecule differs from **1R-Chl** and **1S-Chl** by 79.1 amu. From each sample, a 200  $\mu\text{L}$  serum aliquot was taken to which 2  $\mu\text{L}$  of 100 pmol/ $\mu\text{L}$  control polyamide solution was added along with 800  $\mu\text{L}$  of chilled 100 % methanol. Samples were vigorously vortexed for 1 min and then incubated for 15 min at 4 °C. They were subsequently centrifuged at 12,000 rpm for 10 min, and the supernatant was harvested and concentrated to 50  $\mu\text{L}$  by methanol extraction. Standards with concentrations from 100 to 12,500 fmol/ $\mu\text{L}$  of **1R-Chl** or **1S-Chl** were prepared in serum and extracted by the above procedure. Concentrations were determined by electrospray mass spectrometry with an Agilent 1100 single quadrupole instrument coupled to an Agilent 1100 liquid chromatography system, and a 50 mm x 2.0 mm C18 column for injection and separation.

## Acknowledgements

This work was supported by grants from the National Cancer Institute (CA107311) and the Department of Defense (CM043013) to J.M.G. and P.B.D. M.E.F. is supported by a pre-doctoral NIH NRSA training grant. We thank the staff of The Scripps Research Institute Department of Animal Resources, and The Scripps Mass Spectrometry and DNA Microarray facilities for technical assistance.

## References

1. Lee, S. T.; Neelapu, S. S.; Kwak, L. W. *Yonsei Med. J.* **2007.** 48, 1-10.
2. Hengstler, J. G.; Bockamp, E. O.; Hermes, M.; Brulport, M.; Bauer, A.; Schormann, W.; Schiffer, I. B.; Hausherr, C.; Eshkind, L.; Antunes, C.; Franzen, A.; Krishnamurthi, K.; Lausch, E.; Lessig, R.; Chakrabarti, T.; Prawitt, D.; Zabel, B.; Spangenberg, C. *Curr. Cancer Drug Targets.* **2006.** 6, 603-612.
3. Johnston, S. R. D.; Leary, A. *Drugs Today.* **2006.** 42, 441-453.
4. Ren, R. B. *Curr. Opin. Hematol.* **2004.** 11, 25-34.
5. O'Mahony, D.; Bishop, M. R. *Front. Biosci.* **2006.** 11, 1620-1635.
6. Reid, A.; Vidal, L.; Shaw, H.; de Bono, J. *Eur. J. Cancer.* **2007.** 43, 481-489.
7. O'Hare, T.; Corbin, A. S.; Druker, B. J. *Curr. Opin. Genet. Dev.* **2006.** 16, 92-99.
8. Schittenhelm, M. M.; Shiraga, S.; Schroeder, A.; Corbin, A. S.; Griffith, D.; Lee, F. Y.; Bokemeyer, C.; Deininger, M. W. N.; Druker, B. J.; Heinrich, M. C. *Cancer Res.* **2006.** 66, 473-481.
9. Filippini, C.; Bisiach, M.; Tagliabue, G.; Dincalci, M.; Ubezio, P. *Int. J. Cancer.* **1997.** 72, 801-809.
10. Izbicka, E.; Tolcher, A. W. *Curr. Opin. Investig. Drugs.* **2004.** 5, 587-591.
11. Neidle, S.; Thurston, D. E. *Nat. Rev. Cancer.* **2005.** 5, 285-296.
12. Weiss, G. R.; Poggesi, I.; Rocchetti, M.; DeMaria, D.; Mooneyham, T.; Reilly, D.; Vitek, L. V.; Whaley, F.; Patricia, E.; Von Hoff, D. D.; O'Dwyer, P. *Clin. Cancer Res.* **1998.** 4, 53-59.
13. Lockhart, A. C.; Howard, M.; Hande, K. R.; Roth, B. J.; Berlin, J. D.; Vreeland, F.; Campbell, A.; Fontana, E.; Fiorentini, F.; Fowst, C.; Paty, V. A.; Lankford, O.; Rothenberg, M. L. *Clin. Cancer Res.* **2004.** 10, 468-475.
14. Broggini, M.; Coley, H. M.; Mongelli, N.; Pesenti, E.; Wyatt, M. D.; Hartley, J. A.; Dincalci, M. *Nucleic Acids Res.* **1995.** 23, 81-87.
15. Broggini, M.; Marchini, S.; Fontana, E.; Moneta, D.; Fowst, C.; Geroni, C. *Anti-Cancer Drugs.* **2004.** 15, 1-6.

16. Herzig, M. C. S.; Trevino, A. V.; Arnett, B.; Woynarowski, J. M. *Biochemistry*. **1999**. *38*, 14045-14055.
17. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**. *13*, 284-299.
18. Dudouet, B.; Burnett, R.; Dickinson, L. A.; Wood, M. R.; Melander, C.; Belitsky, J. M.; Edelson, B.; Wurtz, N.; Briehn, C.; Dervan, P. B.; Gottesfeld, J. M. *Chem. Biol.* **2003**. *10*, 859-867.
19. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003**. *100*, 12063-12068.
20. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004**. *32*, 2802-2818.
21. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**. *35*, 363-370.
22. Dickinson, L. A.; Gulizia, R. J.; Trauger, J. W.; Baird, E. E.; Mosier, D. E.; Gottesfeld, J. M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1998**. *95*, 12890-12895.
23. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2004**. *101*, 16768-16773.
24. Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2007**. *104*, 10418-10423.
25. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *ACS Chem. Biol.* **2007**. *2*, 561-571.
26. Janssen, S.; Cuvier, O.; Muller, M.; Laemmli, U. K. *Mol. Cell.* **2000**. *6*, 1013-1024.
27. Janssen, S.; Durussel, T.; Laemmli, U. K. *Mol. Cell.* **2000**. *6*, 999-1011.
28. Burnett, R.; Melander, C.; Puckett, J. W.; Son, L. S.; Wells, R. D.; Dervan, P. B.; Gottesfeld, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **2006**. *103*, 11497-11502.
29. Wurtz, N. R.; Dervan, P. B. *Chem. Biol.* **2000**. *7*, 153-161.
30. Oyoshi, T.; Kawakami, W.; Narita, A.; Bando, T.; Sugiyama, H. *J. Am. Chem. Soc.* **2003**. *125*, 4752-4754.
31. Shinohara, K.; Narita, A.; Oyoshi, T.; Bando, T.; Teraoka, H.; Sugiyama, H. *J. Am. Chem. Soc.* **2004**. *126*, 5113-5118.



32. Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B. S.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M. *Chem. Biol.* **2004**. *11*, 1583-1594.
33. Tsai, S. M.; Farkas, M. E.; Chou, C. J.; Gottesfeld, J. M.; Dervan, P. B. *Nucleic Acids Res.* **2007**. *35*, 307-316.
34. Farkas, M. E.; Tsai, S. M.; Dervan, P. B. *Bioorg. Med. Chem.* **2007**. *15*, 6927-6936.
35. Alvarez, D.; Chou, C. J.; Latella, L.; Zeitlin, S. G.; Ku, S.; Puri, P. L.; Dervan, P. B.; Gottesfeld, J. M. *Cell Cycle*. **2006**. *5*, 1537-1548.
36. Garrity, P. A.; Wold, B. J. *Proc. Natl. Acad. Sci. U. S. A.* **1992**. *89*, 1021-1025.
37. Sabatino, M. A.; Colombo, T.; Geroni, C.; Marchini, S.; Broggini, M. *Clin. Cancer Res.* **2003**. *9*, 5402-5408.
38. Shinohara, K.; Sasaki, S.; Minoshima, M.; Bando, T.; Sugiyama, H. *Nucleic Acids Res.* **2006**. *34*, 1189-1195.
39. Druker, B. J.; Sawyers, C. L.; Kantarjian, H.; Resta, D. J.; Reese, S. F.; Ford, J. M.; Capdeville, R.; Talpaz, M. *N. Engl. J. Med.* **2001**. *344*, 1038-1042.
40. Druker, B. J.; Talpaz, M.; Resta, D. J.; Peng, B.; Buchdunger, E.; Ford, J. M.; Lydon, N. B.; Kantarjian, H.; Capdeville, R.; Ohno-Jones, S.; Sawyers, C. L. *N. Engl. J. Med.* **2001**. *344*, 1031-1037.
41. Jahagirdar, B. N.; Miller, J. S.; Shet, A.; Verfaillie, C. M. *Exp. Hematol.* **2001**. *29*, 543-556.
42. Capdeville, R.; Buchdunger, E.; Zimmermann, J.; Matter, A. *Nat. Rev. Drug Discov.* **2002**. *1*, 493-502.
43. von Bubnoff, N.; Peschel, C.; Duyster, J. *Leukemia*. **2003**. *17*, 829-838.
44. Nimmanapalli, R.; Bhalla, K. *Curr. Opin. Oncol.* **2002**. *14*, 616-620.
45. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**. *118*, 6141-6146.
46. Trauger, J. W.; Dervan, P. B. *Meth. Enzymol.* **2001**. *340*, 450-466.

### **III. Substituent Effects in the $\gamma$ Turn Unit of Hairpin Py-Im Polyamides**

## Chapter IIIA

### Next Generation Hairpin Polyamides with (*R*)-3,4-Diaminobutyric Acid Turn Unit

*The text of this chapter was taken in part from a manuscript co-authored with Christian Dose, David M. Chenoweth, and Peter B. Dervan (California Institute of Technology)*

(Dose, C.; Farkas, M.E.; Chenoweth, D.M.; Dervan, P.B. *J. Am. Chem. Soc.* **2008**, *130*, 6859-6866)

## Abstract

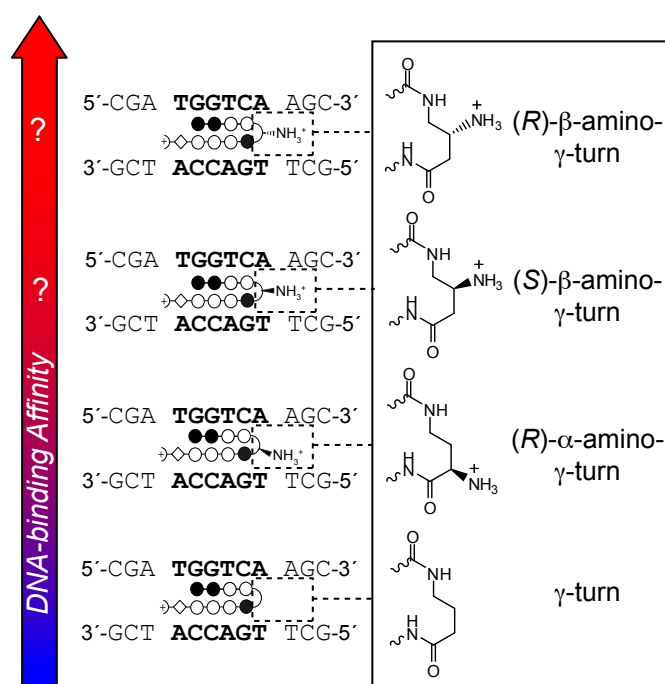
The characterization of a new class of pyrrole-imidazole hairpin polyamides with  $\beta$ -amino- $\gamma$ -turn units for recognition of the DNA minor groove is reported. A library of eight hairpins containing (*R*)- and (*S*)-3,4-diaminobutyric acid ( $\beta$ -amino- $\gamma$ -turn) has been synthesized, and the impact of the molecules on DNA-duplex stabilization was studied for comparison with the parent  $\gamma$ -aminobutyric acid ( $\gamma$ -turn) and standard (*R*)-2,4-diaminobutyric acid ( $\alpha$ -amino- $\gamma$ -turn)-linked eight-ring polyamides. For some, but not all sequence compositions, melting temperature analyses have revealed that both enantiomeric forms of the  $\beta$ -amino- $\gamma$ -turn increase the DNA-binding affinity of polyamides relative to the (*R*)- $\alpha$ -amino- $\gamma$ -turn. The (*R*)- $\beta$ -amine residue may be an attractive alternative for constructing hairpin polyamide conjugates. Biological assays have shown that (*R*)- $\beta$ -amino- $\gamma$ -turn hairpins are able to inhibit androgen receptor-mediated gene expression in cell culture similar to hairpins bearing the standard (*R*)- $\alpha$ -amino- $\gamma$ -turn, from which we infer that they are cell-permeable.

## Introduction

The ability to modulate the expression of eukaryotic gene networks by small molecules is a challenge in the field of chemical biology. Hairpin pyrrole-imidazole polyamides are a class of programmable small molecules that bind to the minor groove of DNA with affinities similar to transcription factors and have been shown to inhibit gene expression in living cells by interfering with transcription factor/DNA interfaces.<sup>1-3</sup> The DNA sequence specificity of polyamides arises from interactions of pairs of the aromatic amino acids *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and *N*-methylhydroxypyrrole (Hp) with the edges of the Watson-Crick base pairs.<sup>4,5</sup> The generality of the polyamide pairing rules has been demonstrated by numerous studies,<sup>6-9</sup> and applications of polyamide conjugates include DNA alkylations,<sup>10-13</sup> DNA templated ligations,<sup>14</sup> sequence-specific DNA intercalators,<sup>15-17</sup> fluorescent DNA paints,<sup>18,19</sup> DNA nanoarchitectures,<sup>20-22</sup> and transcription factor mimics.<sup>23-28</sup> Efforts have been made to improve the DNA-binding properties of hairpin polyamides with modified turn units.<sup>29-32</sup> Substitution of  $\gamma$ -aminobutyric acid ( $\gamma$ -turn) by (*R*)-2,4-diaminobutyric acid ( $\alpha$ -amino- $\gamma$ -turn) increases the DNA-binding affinity by  $\sim 15$ -fold.<sup>30,33</sup> In contrast, hairpins containing the opposite enantiomer, (*S*)- $\alpha$ -amino- $\gamma$ -turn, result in diminished binding affinities. This decrease is most likely caused by an unfavorable steric clash of the amine residue with the DNA minor groove.<sup>30</sup> Sugiyama and co-workers have introduced polyamides containing the  $\alpha$ -hydroxy- $\gamma$ -turn.<sup>31</sup> These hairpins provide discrimination for A·T/T·A base pairs at the turn position, although a  $\sim 70$ -fold reduced DNA-binding affinity relative to analogous (*R*)- $\alpha$ -amino- $\gamma$ -turn-linked polyamides has been observed.

Here we introduce a new class of hairpin polyamides which are linked by 3,4-diaminobutyric acid which results in a  $\beta$ -amine residue at the turn unit ( $\beta$ -amino- $\gamma$ -turn) (Figure IIIA.1). DNA binding affinities of four different eight-ring polyamide core sequences (with incrementally increasing numbers of Im/Py pairs) have been investigated and were compared to analogous hairpins bearing the parent  $\gamma$ -turn and the standard (*R*)-

$\alpha$ -amino- $\gamma$ -turn. We show that for certain series of hairpin polyamides both enantiomers of the  $\beta$ -amino- $\gamma$ -turn are able to increase the relative DNA binding affinity. However, this is sequence context dependent. Biological assays reveal that hairpin polyamides bearing the (*R*)- $\beta$ -amino- $\gamma$ -turn are able to inhibit specific gene expression in cell culture, which is taken as evidence of cell permeability.

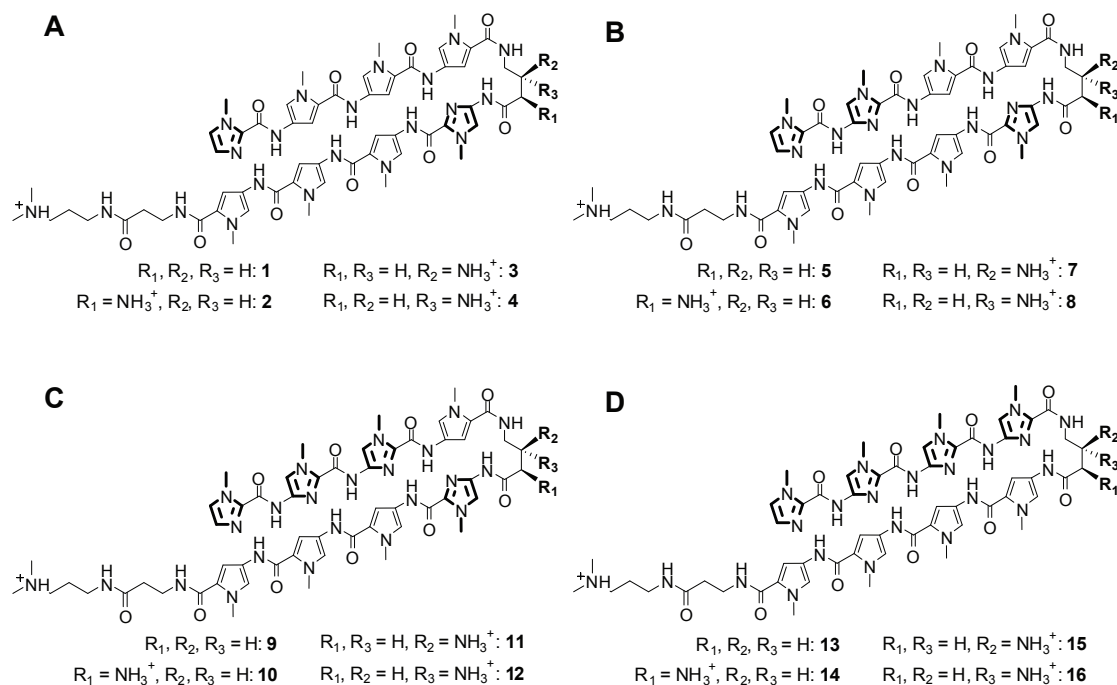


**Figure IIIA.1.** Schematic representation of hairpin polyamides with increased DNA-binding affinity caused by varying  $\gamma$ -turn units. Hairpin polyamides targeted to DNA sequence 5'-TGGTCA-3' are shown as ball-and-stick models. Ball-and-stick representation legend: black and white circles represent *N*-methylimidazole and *N*-methylpyrrole units, respectively, half-circles represent  $\gamma$ -aminobutyric acid, white diamonds represent  $\beta$ -alanine units, and half-circles containing a cross represent 3-(dimethylamino)-1-propylamine (Dp) as tail.

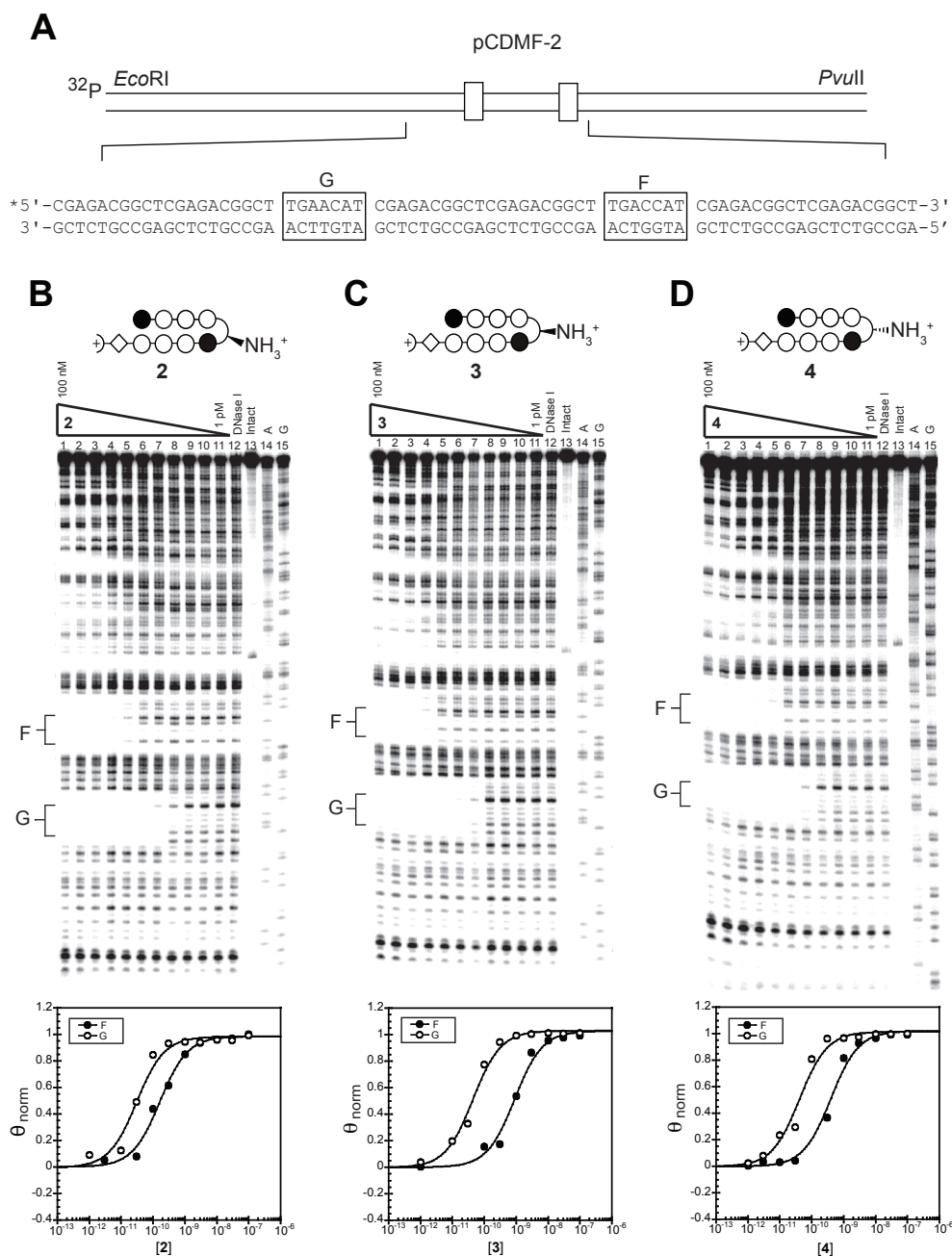
## Results and Discussion

### Thermal stabilization of DNA duplexes by hairpin polyamides

Hairpin polyamides **1-16** were synthesized with different Im/Py and Py/Py compositions targeted to four DNA sequences with increasing G/C content 5'-TGTTCA-3', 5'-TGGTCA-3', 5'-TGGGCA-3', and 5'-TGGGGA-3' (Figure IIIA.2). The energetics of DNA-binding properties of polyamides are typically characterized by quantitative DNase I footprint titrations.<sup>34</sup> These measurements provide precise information regarding the affinity and specificity of DNA/polyamide complexes. Unfortunately, quantitative footprinting experiments revealed similar equilibrium association constants ( $K_a$  values  $\sim 2 \times 10^{10} \text{ M}^{-1}$ ) for hairpins **1-8**, reaching an upper limit of the standard procedure (Figure IIIA.3 - 6, Table IIIA.1). For quantitative footprinting experiments, the DNA concentrations of equilibrium mixtures should be at least 10-fold less than the total association constant of the DNA/ligand complex in order to ensure the approximation  $[\text{ligand}]_{\text{free}} = [\text{ligand}]_{\text{total}}$  for numerical analysis. However, the concentration of the labeled

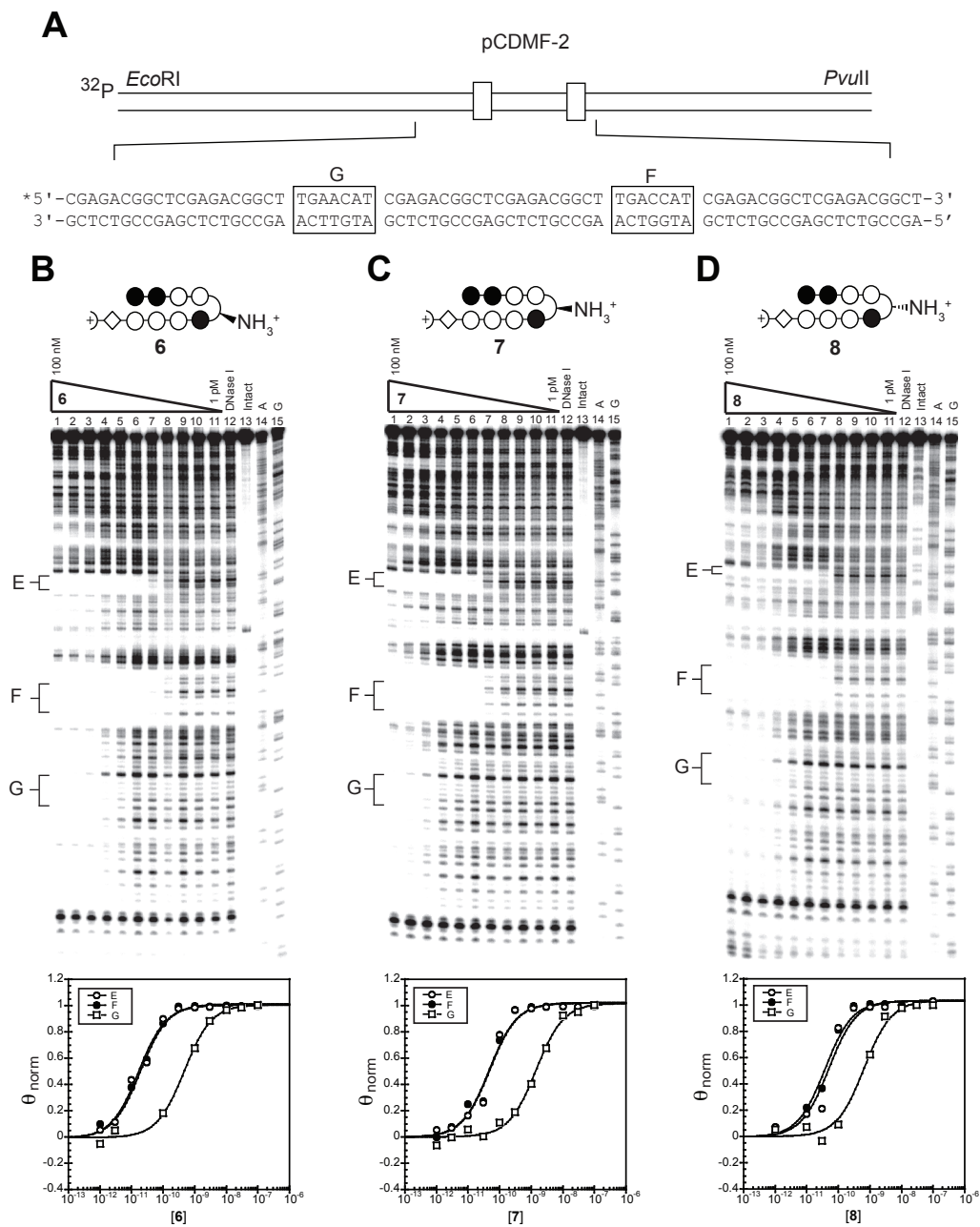


**Figure IIIA.2.** Chemical structures for hairpins **1-16** targeted to DNA sequences: (A) 5'-TGTTCA-3', (B) 5'-TGGTCA-3', (C) 5'-TGGGCA-3', and (D) 5'-TGGGGA-3'.

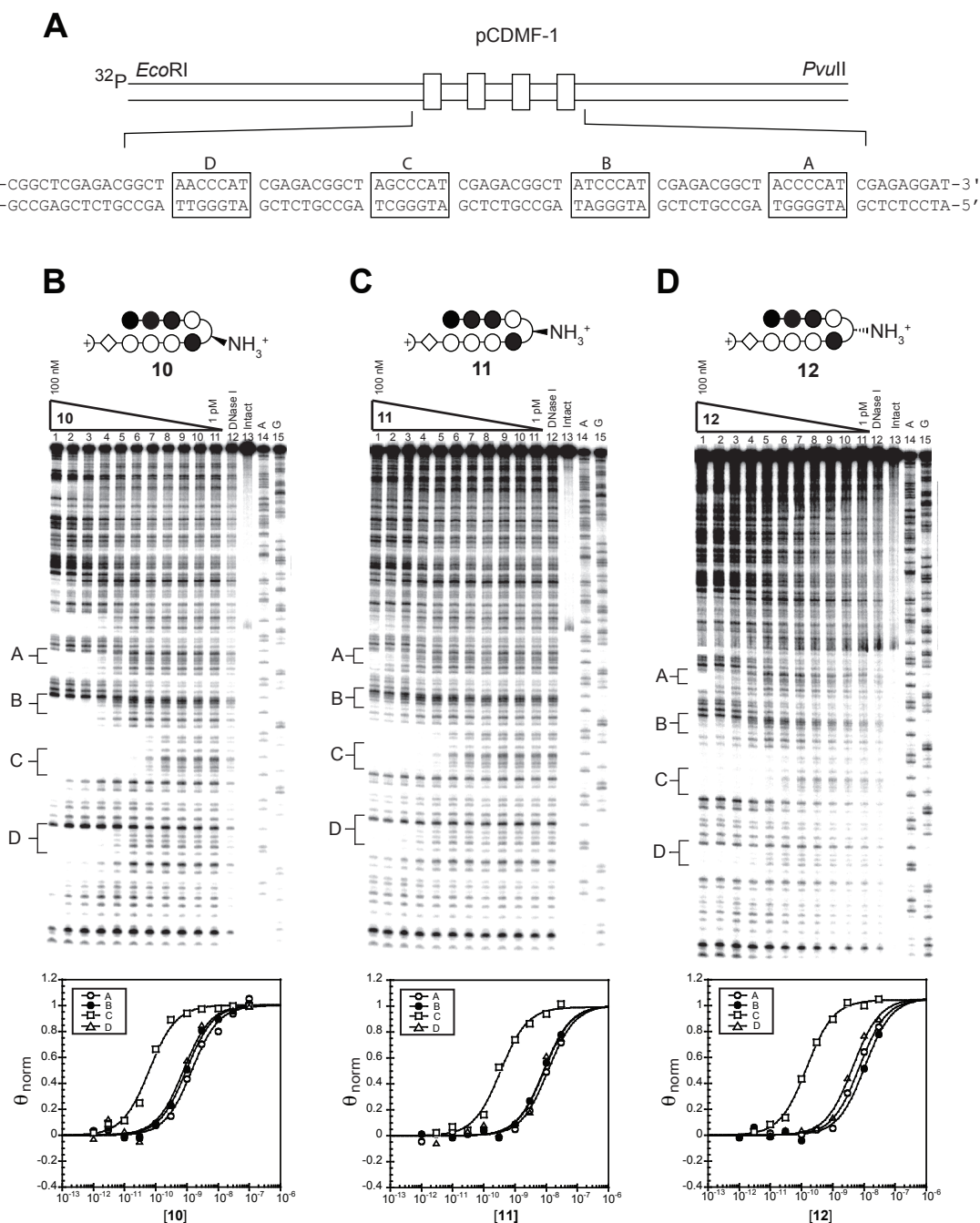


**Figure IIIA.3.** Quantitative DNase I footprint titration experiments on (A) the 285 base pair, 5' end-labeled PCR product of plasmid pCDMF-2 for polyamides (B) **2**, (C) **3**, and (D) **4**: lanes 1-11, 100 nM, 30 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by its respective binding isotherms (below).

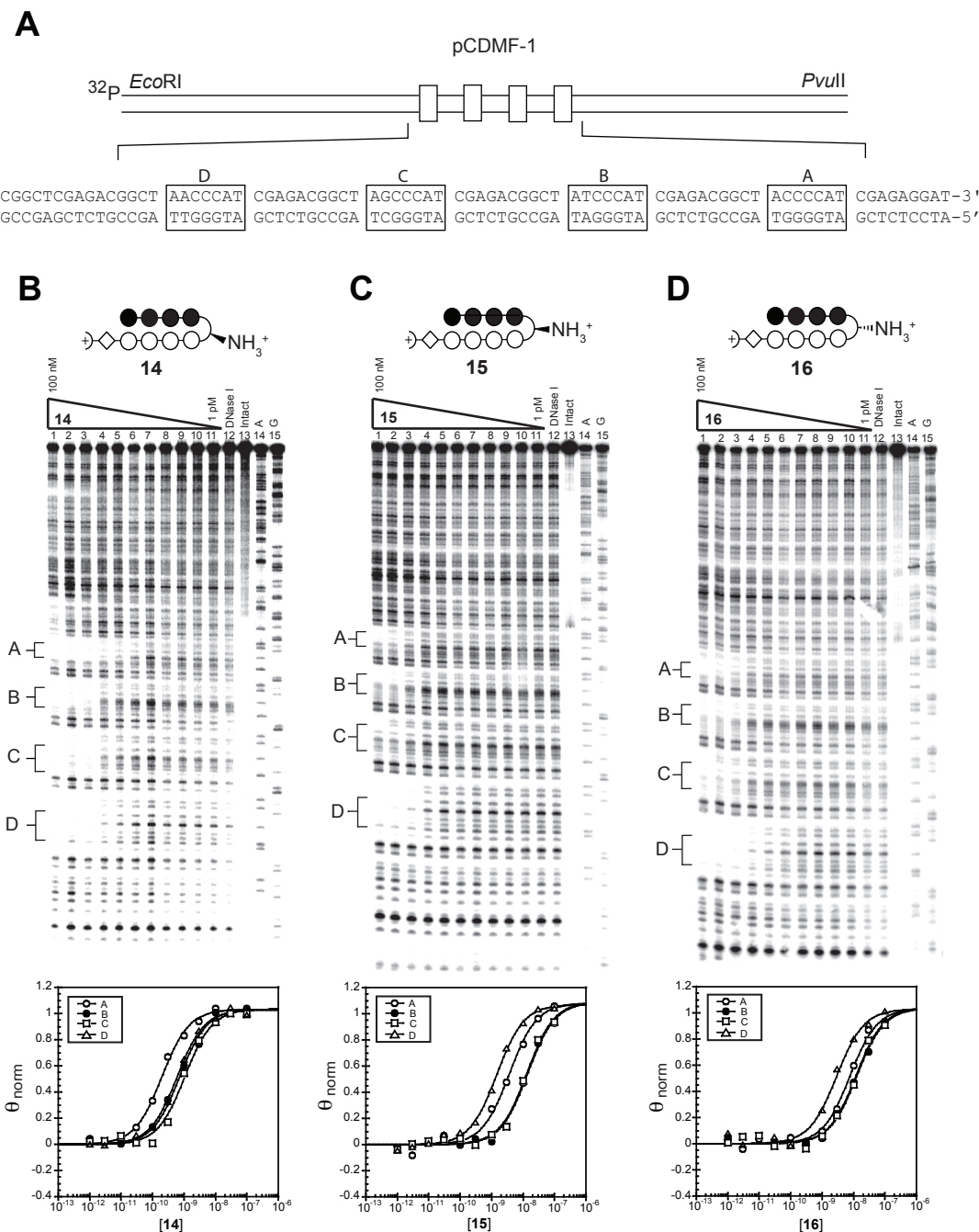




**Figure IIIA.4.** Quantitative DNase I footprint titration experiments on (A) the 285 base pair, 5' end-labeled PCR product of plasmid pCDMF-2 for polyamides (B) **6**, (C) **7**, and (D) **8**: lanes 1-11, 100 nM, 30 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by its respective binding isotherms. Site E is a matched binding site located within the pUC19 plasmid (below).



**Figure IIIA.5.** Quantitative DNase I footprint titration experiments on the (A) 293 base pair, 5' end-labeled PCR product of plasmid pCDMF-1 for polyamides (B) **10**, (C) **11**, and (D) **12**: lanes 1-11, 100 nM, 30 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by its respective binding isotherms (below).



**Figure IIIA.6.** Quantitative DNase I footprint titration experiments on (A) the 293 base pair, 5' end-labeled PCR product of plasmid pCDMF-1 for polyamides (B) **14**, (C) **15**, and (D) **16**: lanes 1-11, 100 nM, 30 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by its respective binding isotherms (below).

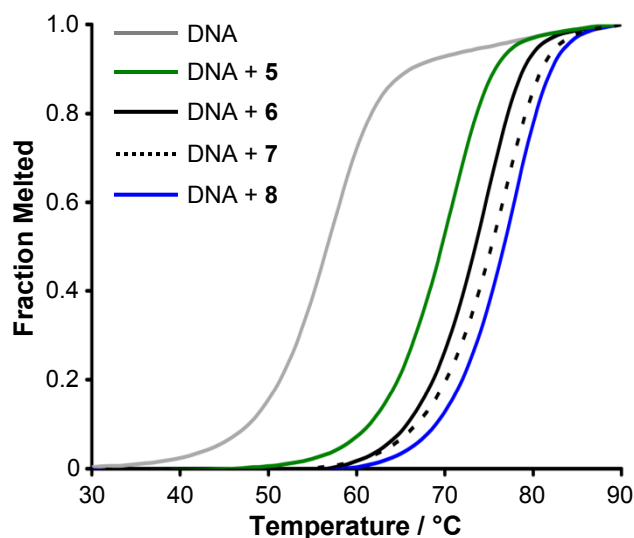
Polyamides		<b>A•T</b>		<b>T•A</b>	
		5'-CGA <b>TGTTCA</b> AGC-3'		5'-CGA <b>TGTTCT</b> AGC-3'	
		$K_a / M^{-1}$		$K_a / M^{-1}$	
	(1)	$3.0 (\pm 0.8) \times 10^{10 [b]}$		n. d.	
	(2)	$2.6 (\pm 0.6) \times 10^{10}$		n. d.	
	(3)	$2.1 (\pm 0.1) \times 10^{10}$		n. d.	
	(4)	$2.7 (\pm 0.3) \times 10^{10}$		n. d.	
		5'-CGA <b>TGGTCA</b> AGC-3'		5'-CGA <b>TGGTCT</b> AGC-3'	
		$K_a / M^{-1}$		$K_a / M^{-1}$	
		$K_a / M^{-1}$		$K_a / M^{-1}$	
	(5)	$1.3 (\pm 0.7) \times 10^{10 [b]}$		n. d.	
	(6)	$3.1 (\pm 0.5) \times 10^{10}$		n. d.	
	(7)	$2.4 (\pm 0.3) \times 10^{10}$		n. d.	
	(8)	$2.3 (\pm 0.3) \times 10^{10}$		n. d.	
		5'-CGA <b>TGGGCA</b> AGC-3'		5'-CGA <b>TGGGCT</b> AGC-3'	
		$K_a / M^{-1}$		$K_a / M^{-1}$	
		$K_a / M^{-1}$		$K_a / M^{-1}$	
	(9)	n. d.		n. d.	
	(10)	n. d.		$1.5 (\pm 0.2) \times 10^{10}$	
	(11)	n. d.		$3.0 (\pm 0.4) \times 10^9$	
	(12)	n. d.		$5.9 (\pm 0.9) \times 10^9$	
		5'-CGA <b>TGGGGA</b> AGC-3'		5'-CGA <b>TGGGGT</b> AGC-3'	
		$K_a / M^{-1}$		$K_a / M^{-1}$	
		$K_a / M^{-1}$		$K_a / M^{-1}$	
	(13)	$2.8 (\pm 0.2) \times 10^{10 [b]}$		n. d.	
	(14)	n. d.		$6.6 (\pm 1.8) \times 10^9$	
	(15)	n. d.		$9.4 (\pm 3.0) \times 10^7$	
	(16)	n. d.		$2.1 (\pm 0.6) \times 10^8$	

**Table IIIA.1.** Equilibrium association constants for hairpin polyamides determined by quantitative DNase I footprint titrations. Equilibrium association constants reported are mean values from at least three quantitative DNase I footprint titration experiments. Standard deviations are shown in parentheses. <sup>[b]</sup>Equilibrium association constants have been reported previously.<sup>33</sup> (n. d. = not determined)

















DNA fragment specified by the standard DNA/polyamide footprinting protocol is  $\sim 5$  pM.<sup>34</sup> Consequently, DNA association constants become compressed and hence unreliable for comparison studies at  $K_a$  values  $\geq 2 \times 10^{10} M^{-1}$ .<sup>35,36</sup>

Prior results have shown that the increase in melting temperature ( $\Delta T_m$ ) of DNA

duplexes bound by hairpin polyamides correlates with DNA-binding affinity and can be utilized to detect single base pair mismatched DNA/polyamide complexes.<sup>37,38</sup> Accordingly, we have used melting temperature analysis for dissecting differences in DNA affinities of hairpin polyamides. Spectroscopic analyses were performed on 12mer DNA duplexes containing the appropriate match sequence in the absence and presence of polyamides in order to derive the desired  $\Delta T_m$  values (Figure IIIA.7). Table IIIA.2 shows that all hairpins provided an increase in melting temperature relative to the individual DNA duplexes, confirming the formation of DNA/polyamide complexes. As expected, spectroscopic analysis with (*R*)- $\alpha$ -amino- $\gamma$ -turn hairpins revealed stronger stabilizations than the parent  $\gamma$ -turn analogues. For example, achiral polyamide **1** targeted to DNA sequence 5'-TGTTCA-3' resulted in a  $\Delta T_m$  value of 15.9 °C, while chiral hairpin (*R*)- $\alpha$ -**2** led to a 3.6 °C higher melting temperature ( $\Delta T_m$  = 19.5 °C). Remarkably, melting temperature analyses in the presence of  $\beta$ -amino- $\gamma$ -turn hairpins (*S*)- $\beta$ -**3** ( $\Delta T_m$  = 20.9 °C) and (*R*)- $\beta$ -**4** ( $\Delta T_m$  = 22.2 °C) revealed higher  $\Delta T_m$  values compared to those for the  $\alpha$ -series hairpin (*R*)- $\alpha$ -**2** ( $\Delta T_m$  = 19.5 °C).



**Figure IIIA.7.** Normalized UV denaturation profiles of 12mer DNA duplex 5'-CGATGGTCAAGC-3'/5'-GCTTGACCATCG-3' in the absence and presence of hairpin polyamides **5-8**.

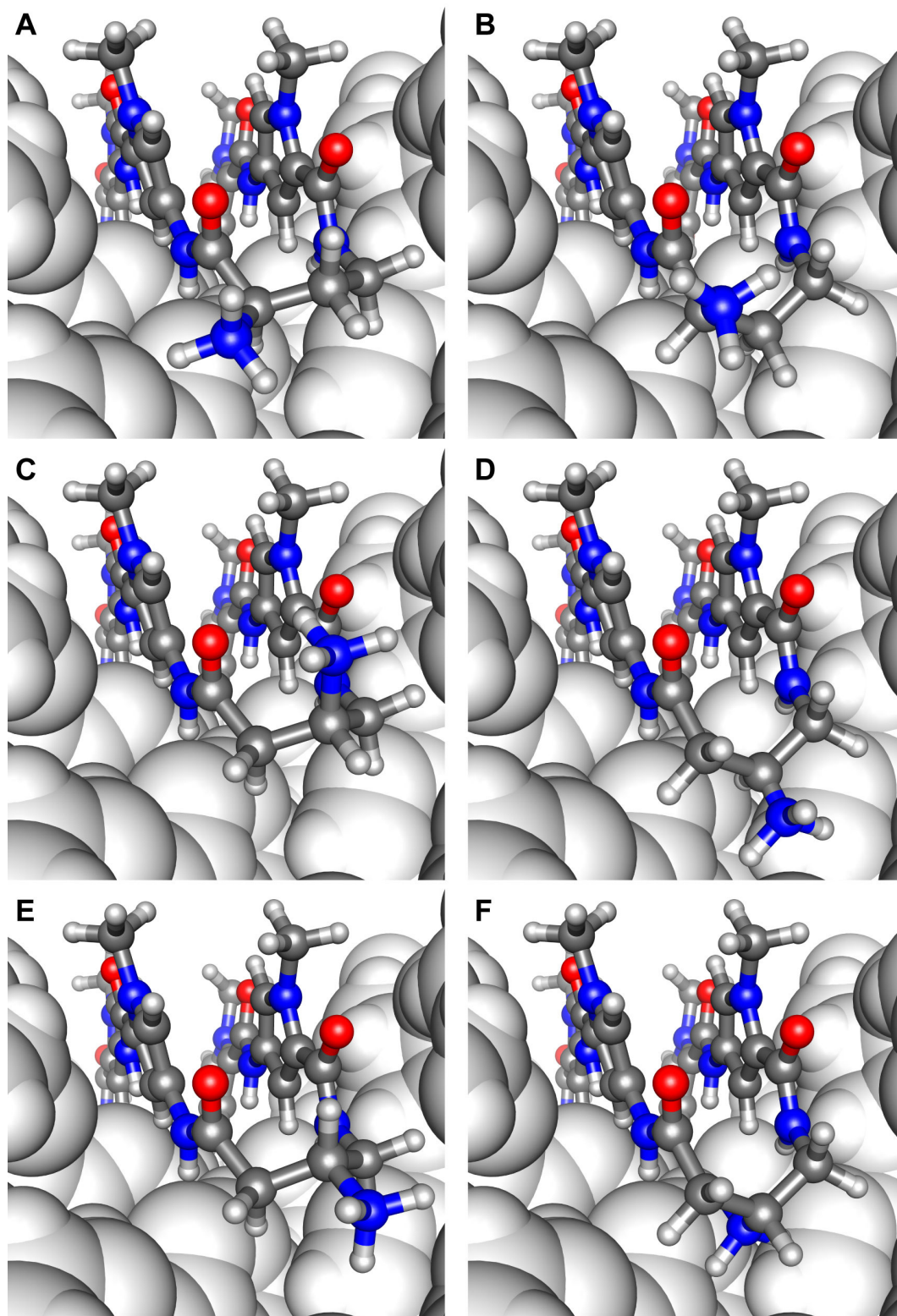
Polyamides		<b>A•T</b>		<b>T•A</b>	
		5'-CGA <b>TGTTCA</b> AGC-3'		5'-CGA <b>TGTTCT</b> AGC-3'	
		$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$
—		54.0 ( $\pm 0.2$ )	—	n.d.	
	(1)	69.9 ( $\pm 0.3$ )	15.9	n.d.	
	(2)	73.5 ( $\pm 0.2$ )	19.5	n.d.	
	(3)	74.9 ( $\pm 0.2$ )	20.9	n.d.	
	(4)	76.2 ( $\pm 0.2$ )	22.2	n.d.	
		5'-CGA <b>TGGTCA</b> AGC-3'		5'-CGA <b>TGGTCT</b> AGC-3'	
—		57.2 ( $\pm 0.1$ )	—	55.8 ( $\pm 0.1$ )	—
	(5)	70.6 ( $\pm 0.2$ )	13.4	69.0 ( $\pm 0.3$ )	13.2
	(6)	74.1 ( $\pm 0.3$ )	16.9	72.9 ( $\pm 0.2$ )	17.1
	(7)	76.1 ( $\pm 0.2$ )	18.9	73.2 ( $\pm 0.1$ )	17.4
	(8)	77.5 ( $\pm 0.3$ )	20.3	74.2 ( $\pm 0.1$ )	18.4
		5'-CGA <b>TGGGCA</b> AGC-3'		5'-CGA <b>TGGGCT</b> AGC-3'	
—		60.2 ( $\pm 0.2$ )	—	59.8 ( $\pm 0.3$ )	—
	(9)	68.8 ( $\pm 0.2$ )	8.6	67.4 ( $\pm 0.3$ )	7.6
	(10)	73.4 ( $\pm 0.2$ )	13.2	72.0 ( $\pm 0.1$ )	12.2
	(11)	73.5 ( $\pm 0.1$ )	13.3	70.5 ( $\pm 0.3$ )	10.7
	(12)	73.8 ( $\pm 0.1$ )	13.6	71.3 ( $\pm 0.3$ )	11.5
		5'-CGA <b>TGGGGA</b> AGC-3'		5'-CGA <b>TGGGGT</b> AGC-3'	
—		57.5 ( $\pm 0.1$ )	—	57.9 ( $\pm 0.1$ )	—
	(13)	60.9 ( $\pm 0.1$ )	3.4	61.4 ( $\pm 0.3$ )	3.5
	(14)	66.6 ( $\pm 0.1$ )	9.1	67.0 ( $\pm 0.1$ )	9.1
	(15)	64.2 ( $\pm 0.1$ )	6.7	64.2 ( $\pm 0.1$ )	6.3
	(16)	64.3 ( $\pm 0.3$ )	6.8	64.4 ( $\pm 0.3$ )	6.5

**Table IIIA.2.** Melting temperatures of DNA/polyamide complexes for A•T and T•A base pairs at the turn position of hairpin polyamides. All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses (n.d. = not determined).  $\Delta T_m$  values are given as  $T_m^{(\text{DNA/polyamide})} - T_m^{(\text{DNA})}$ .

The same trend was observed for hairpins **5-8** targeted to DNA sequence 5'-TGGTCA-3' (Table IIIA.2, Figure IIIA.2). First, it is noteworthy that both the (*R*)- and (*S*)- $\beta$ -amino- $\gamma$ -turn generated higher melting temperatures than the standard (*R*)- $\alpha$ -amino- $\gamma$ -turn. Second, the enhancement observed for the (*R*)- $\beta$ -series is almost twice that of the (*R*)- $\alpha$ -series (relative to achiral hairpins) targeted to DNA sequences 5'-TGTTCA-3' and 5'-TGGTCA-3'. For example, polyamide (*R*)- $\beta$ -**8** provided a  $\Delta\Delta T_m$  value of 6.9 °C, while the  $\alpha$ -series (*R*)- $\alpha$ -**6** led to a  $\Delta\Delta T_m$  value of 3.5 °C relative to achiral hairpin **5**. Interestingly, by further increasing the numbers of Im/Py pairs in the polyamides, significantly diminished DNA duplex stabilizations have been observed. For example, achiral polyamide **9** and chiral hairpin (*R*)- $\alpha$ -**10** targeted to DNA sequence 5'-TGGGCA-3' yielded  $\Delta T_m$  values of 8.6 and 13.2 °C, while the  $\beta$ -series (*S*)- $\beta$ -**11** and (*R*)- $\beta$ -**12** led to  $\Delta T_m$  values of 13.3 and 13.6 °C, respectively (Table IIIA.2). Even lower melting temperatures were observed for hairpins **13-16** designed to bind DNA sequence 5'-TGGGGA-3'. Both  $\beta$ -amino- $\gamma$ -turns, as in (*S*)- $\beta$ -**15** ( $\Delta T_m = 6.7$ ) and (*R*)- $\beta$ -**16** ( $\Delta T_m = 6.8$ ), resulted in significantly lower  $\Delta T_m$  values than the  $\alpha$ -series analogue (*R*)- $\alpha$ -**14** ( $\Delta T_m = 9.1$ ). These results imply that the impact of polyamide turn units on DNA-duplex stabilization is sequence context dependent.

The general increase in DNA-binding affinity for polyamides containing the (*R*)- $\alpha$ -substituted  $\gamma$ -turn, relative to achiral hairpins, is most likely caused by a superposition of favorable non-covalent interactions of the positively charged substituent and conformational preferences of the turn unit.<sup>30</sup> The (*R*)- $\alpha$ -amino- $\gamma$ -turn can exist in two different conformations. One conformation orients the  $\alpha$ -ammonium in a pseudo-equatorial position (Figure IIIA.8A), which directs the substituent toward the wall of the minor groove with the potential of steric interactions. The alternate conformation places the  $\alpha$ -amine residue in a pseudo-axial position out of the minor groove, orienting the  $\beta$ -methylene to the floor of the double helix (Figure IIIA.8B). Modeling of the (*S*)- $\beta$ -amino- $\gamma$ -turn conformations suggests that the  $\beta$ -ammonium in a pseudo-axial position is directed out of the minor groove, (Figure IIIA.8C) relieving the potential steric interactions with the wall in comparison to





**Figure IIIA.8.** Illustrative models of different turn conformations for hairpin polyamides containing: (A) and (B) the  $(R)$ - $\alpha$ -amino- $\gamma$ -turn; (C) and (D) the  $(S)$ - $\beta$ -amino- $\gamma$ -turn; and (E) and (F) the  $(R)$ - $\beta$ -amino- $\gamma$ -turn bound to the minor groove of DNA (dark gray = carbons, white = hydrogen, blue = nitrogen, red = oxygen).



the  $\alpha$ -series. In contrast, the (*S*)- $\beta$ -amine in a pseudo-equatorial orientation follows the curvature of the minor groove (Figure IIIA.8D). The possibility of favorable non-covalent interactions should exist in both conformations without the detriment of steric interactions. As shown in Figure IIIA.8E, the pseudo-equatorial  $\beta$ -amine residue of the (*R*)- $\beta$ -amino- $\gamma$ -turn is well accommodated in the DNA minor groove, while the pseudo-axial position should result in a steric clash of the substituent with the groove floor (Figure IIIA.8F).

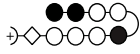
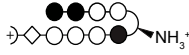
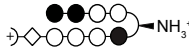

Previous results have shown that polyamides constructed with several continuous Im/Py pairs are over-curved with respect to the DNA minor groove, significantly influencing the DNA binding affinity and sequence specificity.<sup>39,40</sup> We assume that this curvature affects the alignment of the turn units in the DNA minor groove. This is supported by the observation that the presence of fewer continuous imidazoles improves the DNA affinity of  $\beta$ -amino- $\gamma$ -turns while diminishing the affinity of  $\alpha$ -amino- $\gamma$ -turns, and vice-versa. However, illustrative modeling is not sufficient to explain the sequence context dependence of chiral hairpin polyamides, highlighting the pressing need for high resolution structural studies.

### Sequence specificity at the turn position

Hairpin polyamides containing the  $\gamma$ -turn have been shown to possess an equal preference for A·T/T·A over G·C/C·G base pairs at the turn position, presumably for steric reasons.<sup>41</sup> Sugiyama's  $\alpha$ -hydroxy- $\gamma$ -turns have been demonstrated to discriminate A·T versus T·A at the turn position.<sup>31</sup> In order to study the sequence specificity of polyamides **5-16**, we performed melting temperature analyses in the presence of DNA duplexes bearing a T·A base pair at the turn position. Experiments involving hairpins **1-4** have been omitted due to the palindromic core sequence specified by the polyamides. As shown in Table IIIA.2, most  $\gamma$ -turn and (*R*)- $\alpha$ -amino- $\gamma$ -turn hairpins provided similar  $\Delta T_m$  values for T·A and A·T base pairs. In contrast, significantly lower thermal stabilizations for T·A over A·T base pairs were observed for  $\beta$ -amino- $\gamma$ -turn-linked polyamides targeting DNA

sequences 5'-TGGTCA-3' (**5-8**) and 5'-TGGGCA-3' (**9-12**). Even more diminished duplex stabilizations were observed in presence of C•G and G•C base pairs (Table IIIA.3).

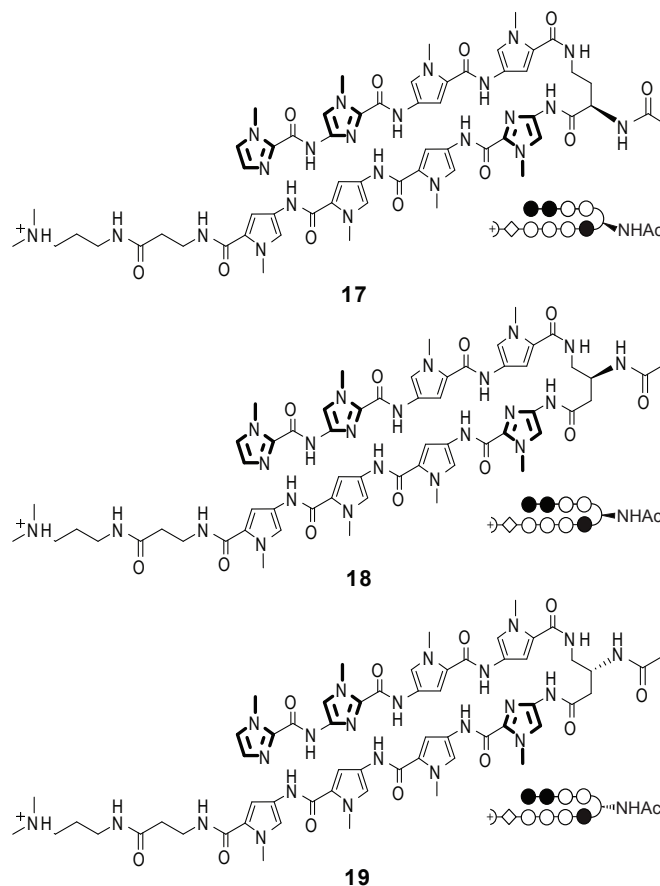
These observations suggest that polyamides containing  $\beta$ -amino- $\gamma$ -turns prefer A•T > T•A >> C•G > G•C base pairs at the turn position. However, sequence specificity studies by using thermal denaturation measurements require binding enthalpies ( $\Delta H_b$ ) of DNA/polyamide complexes in order to determine equilibrium association constants.<sup>37</sup> One could also imagine using six-ring hairpin polyamides with lower DNA binding affinities in order to discriminate sequence specificities at the turn position by quantitative DNase I footprint titration methods.

Polyamides	C•G		G•C	
	5'-CGA <b>TGGTCC</b> AGC-3'		5'-CGA <b>TGGTCG</b> AGC-3'	
	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$
—	59.7 ( $\pm 0.3$ )	—	60.4 ( $\pm 0.2$ )	—
 ( <b>5</b> )	65.9 ( $\pm 0.3$ )	6.2	64.3 ( $\pm 0.1$ )	3.9
 ( <b>6</b> )	67.3 ( $\pm 0.2$ )	7.6	64.7 ( $\pm 0.2$ )	4.3
 ( <b>7</b> )	69.7 ( $\pm 0.3$ )	10.0	66.1 ( $\pm 0.2$ )	5.7
 ( <b>8</b> )	70.1 ( $\pm 0.1$ )	10.4	66.8 ( $\pm 0.2$ )	6.4




**Table IIIA.3.** Melting temperatures of DNA/polyamide complexes for polyamides **5-8** with C•G and G•C base pairs at the turn position DNA binding sites. All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses.  $\Delta T_m$ -values are given as  $T_m^{(\text{DNA/polyamide})} - T_m^{(\text{DNA})}$ .

### Acetylated chiral hairpin polyamides

Several approaches have been reported wherein the (*R*)- $\alpha$ -amino- $\gamma$ -turn was utilized as a position for synthetic modifications of hairpin polyamides.<sup>10-14,42-45</sup> It has been shown that acetylation of the (*R*)- $\alpha$ -amine in six-ring hairpin polyamides results in ~15-fold reduced DNA-binding affinity.<sup>30</sup> To study the tolerance of synthetic modifications in eight-ring polyamides containing the  $\beta$ -amino- $\gamma$ -turns, we examined acetylated hairpins **17-19** by melting temperature analysis (Figure IIIA.9). Indeed, hairpin **17** containing the acetylated (*R*)- $\alpha$ -amino- $\gamma$ -turn yielded a markedly lower  $\Delta T_m$  value (12.8 °C) than non-acetylated analogue **6** ( $\Delta T_m$  = 16.9 °C, Table IIIA.2 and 4). Even more pronounced was the decrease in DNA duplex stabilization for acetylated (*S*)- $\beta$ -amino- $\gamma$ -turn hairpin **18** leading to a  $\Delta T_m$  value of 11.7 °C. Remarkably, the opposite enantiomer (*R*)- $\beta$ -**19** resulted in significantly less destabilization ( $\Delta T_m$  = 17.8 °C). All hairpins lose the positive charge at the turn unit by acetylation. This implies that the cationic state of the amine residue is not the only contribution impacting the energetics of the DNA/polyamide complexes, as evidenced by the differences in melting temperatures between hairpins **17-19**. Increased steric demands of the acetylated substituents may also be responsible for differing binding affinities, due to the restricting DNA minor groove and alternate conformations of the  $\gamma$ -turn units (Figure IIIA.8).



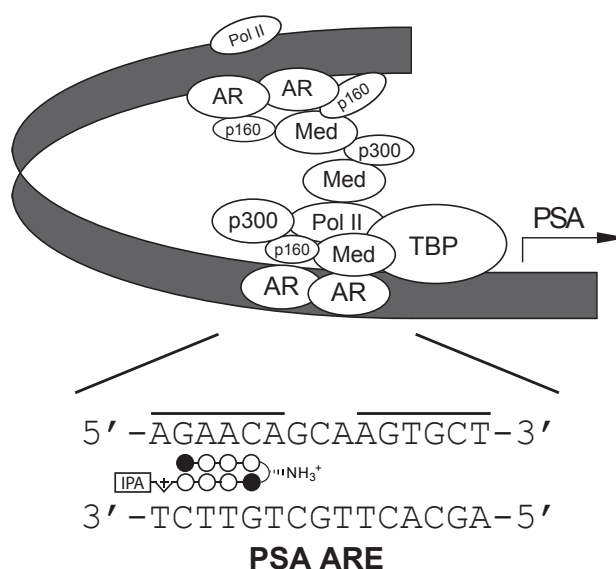
**Figure IIIA.9.** Chemical structures and ball-and-stick models of acetylated hairpin polyamides **17-19** targeted to DNA sequence 5'-TGGTCA-3'.

DNA sequence = 5'-CGA <b>TGGTCA</b> AGC-3'		
Polyamides	$T_m$ / °C	$\Delta T_m$ / °C
—	57.2 ( $\pm 0.2$ )	—
 <b>(17)</b>	70.0 ( $\pm 0.1$ )	12.8
 <b>(18)</b>	68.9 ( $\pm 0.2$ )	11.7
 <b>(19)</b>	75.0 ( $\pm 0.1$ )	17.8

**Table IIIA.4.** Melting temperatures for DNA complexes containing acetylated hairpin polyamides targeted to DNA Sequence 5'-TGGTCA-3'. All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses.  $\Delta T_m$  values are given as  $T_m^{(DNA/polyamide)} - T_m^{(DNA)}$ .

### Biological assay for cell permeability

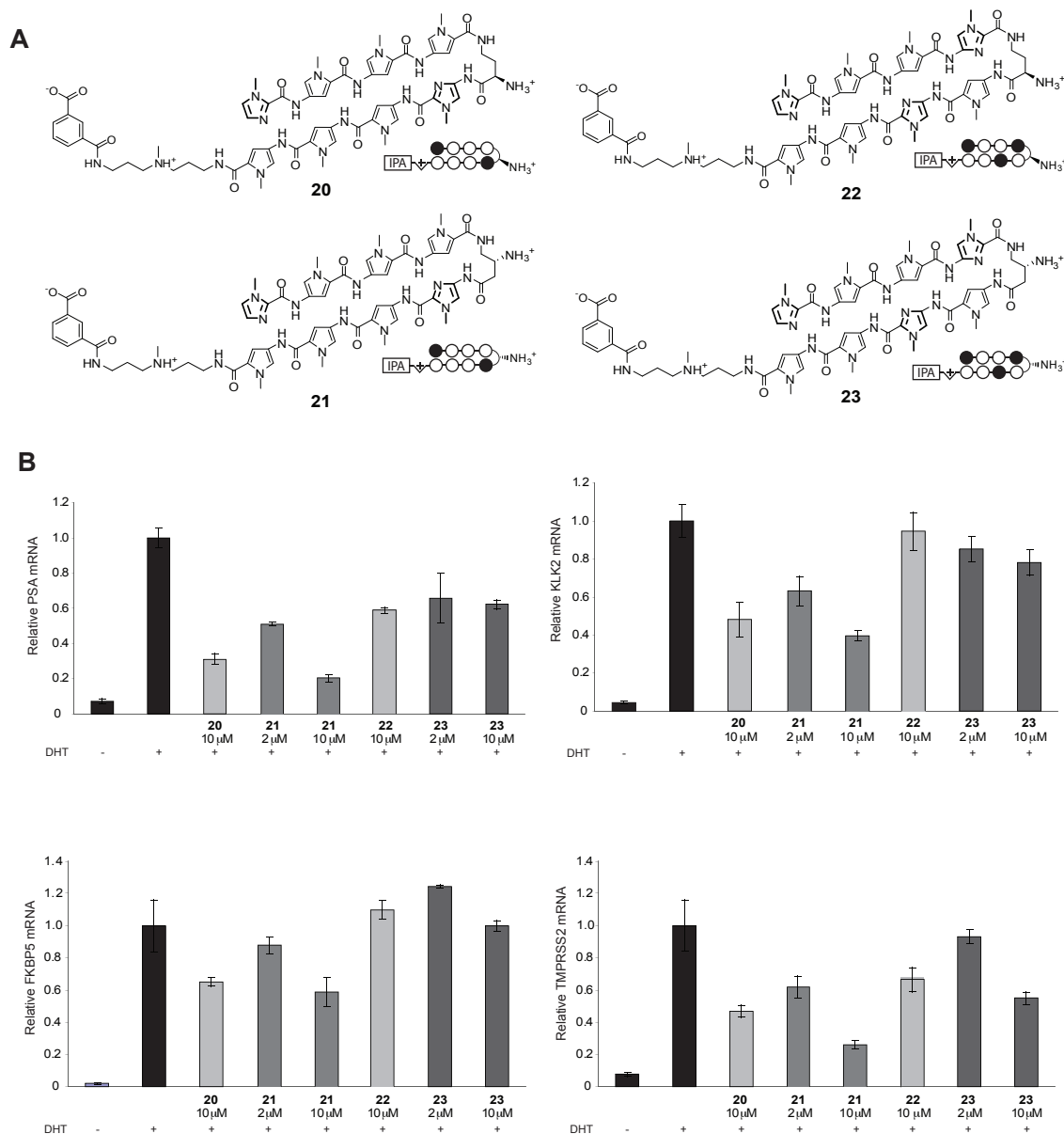
Hairpin polyamide conjugates bearing the standard (*R*)- $\alpha$ -amino- $\gamma$ -turn have been shown to modulate the expression of certain gene pathways in living cells by interfering with transcription factor/DNA interfaces.<sup>1-3</sup> Recently, a hairpin designed to bind the DNA sequence 5'-AGAACA-3', found in the androgen response element (ARE), has been demonstrated to inhibit androgen receptor-mediated expression of prostate specific antigen (PSA) in LNCaP cells (Figure IIIA.10).<sup>2</sup> We utilized this cell culture transcription assay to investigate the cell permeability of (*R*)- $\beta$ -amino- $\gamma$ -turn hairpins because small structural changes within polyamides can influence nuclear uptake properties.<sup>46,47</sup> Hairpin







**Figure IIIA.10.** Schematic representation of the androgen receptor (AR)-mediated transcription complex with the prostate specific antigen (PSA) androgen response element (ARE) DNA sequence.

polyamide **21** was examined in comparison with the previously used (*R*)- $\alpha$ -amino- $\gamma$  turn hairpin **20** (Figure IIIA.11A). Chiral polyamides **22** and **23**, designed to target different DNA sequences, have been used as controls. Melting temperature analyses for polyamide conjugates **20-23** confirmed the results obtained for hairpins **1-4**, revealing highest DNA-duplex stabilizations for (*R*)- $\beta$ -amino- $\gamma$ -turn hairpins (Table IIIA.5). The induction of

PSA mRNA by dihydrotestosterone (DHT) in LNCaP cells treated with matched and mismatched polyamides **20-23** was measured by quantitative real-time RT-PCR. Hairpin **21** provided significant inhibition of AR-mediated expression of PSA mRNA, KLK2, FKBP5, and TMPRSS2 mRNA, which supports cell permeable properties for (*R*)- $\beta$ -amino- $\gamma$ -turn hairpins (Figure IIIA.11B).



**Figure IIIA.11** (A) Chemical structures and ball-and-stick models of matched and mismatched polyamides **20-23**, targeted to DNA sequence 5'-AGAACA-3'. (B) Inhibition of DHT-induced PSA, KLK2, FKBP5, and TMPRSS2 expression by **20-23** measured by quantitative real-time RT-PCR.

DNA sequence = 5'-TTGC <b>AGAACA</b> GCAA-3'		
Polyamides	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$
—	60.1 ( $\pm 0.2$ )	—
 (20)	74.4 ( $\pm 0.2$ )	14.3
 (21)	76.3 ( $\pm 0.2$ )	16.2
 (22)	64.6 ( $\pm 0.1$ )	4.5
 (23)	66.9 ( $\pm 0.1$ )	6.8

**Table IIIA.5.** Melting temperatures of polyamides targeted to DNA-sequence 5'-AGAACA-3' in complex with DNA. All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses.  $\Delta T_m$  values are given as  $T_m^{(\text{DNA/polyamide})} - T_m^{(\text{DNA})}$ .

## Conclusion

Herein we have introduced (*R*)- and (*S*)- $\beta$ -amino- $\gamma$ -turn hairpin polyamides. Eight new polyamides targeting different DNA-binding motifs have been synthesized, and their impact on DNA duplex stabilization in relation to hairpins containing the parent  $\gamma$ -turn and the standard (*R*)- $\alpha$ -amino- $\gamma$ -turn was investigated. It was found that changing the turn unit from the (*R*)- $\alpha$ -amino- $\gamma$ -turn to either enantiomeric forms of the  $\beta$ -amino- $\gamma$ -turn increases the relative DNA-binding affinity of polyamides targeted to 5'-TGTTCA-3' and 5'-TGGTCA-3' but not to 5'-TGGGCA-3' and 5'-TGGGGA-3' sequences, rendering the impact of  $\beta$ -amino-substituted  $\gamma$ -turns sequence context dependent. Acetylation of the (*S*)- $\beta$ -amino- $\gamma$ -turn has been demonstrated to significantly impact the DNA-binding affinity but has minimal effect for the (*R*)- $\beta$ -amino- $\gamma$ -turn, which makes the (*R*)- $\beta$ -amino residue attractive for synthetic modifications and conjugate design. Upper limits presented by DNase I footprinting titrations of high affinity binders rendered melting temperature analysis a more practical choice for dissecting improvements in new turn units of hairpin polyamides. Due to the strong thermal stabilizations reported for eight-ring hairpin polyamides **1-8** targeted to 5'-TGTTCA-3' and 5'-TGGTCA-3' sequences, it is not unreasonable to estimate that the DNA-binding equilibrium association constants are markedly higher than  $K_a = 2 \times 10^{10} \text{ M}^{-1}$ . Biological experiments have demonstrated that (*R*)- $\beta$ -amino- $\gamma$ -turn hairpins possess biological activity to inhibit AR-mediated gene expression within a human cancer cell line and may have uptake properties similar to polyamides bearing the standard (*R*)- $\alpha$ -amino- $\gamma$ -turn. Ongoing work is focused on the use of the next generation hairpins in biological investigations as well as turn unit sequence specificity and high-resolution crystallographic studies for DNA/chiral hairpin polyamide complexes. These efforts will be reported in due course.



## Materials and Methods

### General

Chemicals and solvents were purchased from Sigma-Aldrich and were used without further purification. Boc- $\gamma$ -Abu-OH was purchased from Novabiochem. (*R*)-2,4-Fmoc Dbu(Boc)-OH and Boc- $\beta$ -Ala-PAM resin were purchased from Peptides International. (*R*)-3,4-Cbz-Dbu(Boc)-OH and (*S*)-3,4-Cbz-Dbu(Boc)-OH were purchased from Senn Chemicals AG. All DNA oligomers were purchased HPLC purified from Integrated DNA Technologies. Water (18 M $\Omega$ ) was purified using a Millipore MilliQ purification system. The pH of buffers was adjusted using a Beckman 340 pH/temp meter. Analytical HPLC was performed on a Beckman Gold system equipped with a diode array detector using a Phenomenex Gemini column (5  $\mu$ m particle size, C18 110A, 250  $\times$  4.6 mm, 5  $\mu$ m). Preparative HPLC was performed on a Beckman Gold system equipped with a single-wavelength detector monitoring at 310 nm using a Phenomenex Gemini column (5  $\mu$ m particle size, C18 110A, 250  $\times$  21.2 mm, 5  $\mu$ m). For both analytical and preparative HPLC, solvent A was 0.1% (v/v) aqueous trifluoroacetic acid (TFA) and solvent B was acetonitrile. Solvent gradients were adjusted as needed. Polyamide concentrations were measured in 0.1% (v/v) aqueous TFA on a Hewlett-Packard diode array spectrophotometer “Model 8452 A” and were determined by using an extinction coefficient of 69200 M<sup>-1</sup> · cm<sup>-1</sup> at  $\lambda_{\text{max}}$  near 310 nm. Matrix-assisted, LASER desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on an Applied Biosystems Voyager DR Pro spectrometer using  $\alpha$ -cyano-4-hydroxycinnamic acid as matrix.

### Synthesis of polyamides

Polyamide monomers and oligomers were synthesized as described previously.<sup>48</sup> All  $\beta$ -amino- $\gamma$ -turn hairpins were synthesized by performing the following procedure: the polyamide was cleaved from the resin with 3-(dimethylamino)-1-propylamine, purified by preparative HPLC, and characterized by MALDI-TOF MS, UV-Vis spectroscopy,

and analytical HPLC. A 500 nmol fraction of the Cbz-protected hairpin polyamide was dissolved in a 9:1 mixture (500  $\mu$ L) of TFA and trifluoromethanesulfonic acid (TFMSA). After 5 min reaction time, the solution was flash-frozen by liquid N<sub>2</sub> and overlaid with *N,N'*-dimethylformamide (1 mL). The thawed solution was diluted with 20% aqueous acetonitrile (8 mL), purified by preparative HPLC, and characterized by MALDI-TOF MS, UV-Vis spectroscopy, and analytical HPLC. Acetylated polyamides **17-19** were synthesized by performing the following procedure: A 500 nmol fraction of the polyamide was dissolved in *N,N'*-dimethylformamide (900  $\mu$ L) and a 9:1 mixture of pyridine:acetic anhydride (100  $\mu$ L) was added. After 5 min reaction time, the solution was diluted with 10% aqueous TFA (8 mL), purified by preparative HPLC, and characterized by MALDI-TOF MS, UV-vis spectroscopy, and analytical HPLC. Polyamide conjugates **20-23** were synthesized as described previously.<sup>2</sup>

Polyamide **1**: MALDI-TOF MS C<sub>58</sub>H<sub>72</sub>N<sub>21</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1222.6; found: 1222.7

Polyamide **2**: MALDI-TOF MS C<sub>58</sub>H<sub>73</sub>N<sub>22</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1237.6; found: 1237.8

Polyamide **3**: MALDI-TOF MS C<sub>58</sub>H<sub>73</sub>N<sub>22</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1237.6; found: 1237.8

Polyamide **4**: MALDI-TOF MS C<sub>58</sub>H<sub>73</sub>N<sub>22</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1237.6; found: 1237.8

Polyamide **5**: MALDI-TOF MS C<sub>57</sub>H<sub>71</sub>N<sub>22</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1223.6; found: 1223.5

Polyamide **6**: MALDI-TOF MS C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1238.6; found: 1238.6

Polyamide **7**: MALDI-TOF MS C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1238.6; found: 1238.5

Polyamide **8**: MALDI-TOF MS C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1238.6; found: 1238.5

Polyamide **9**: MALDI-TOF MS C<sub>56</sub>H<sub>70</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1224.6; found: 1224.8

Polyamide **10**: MALDI-TOF MS C<sub>56</sub>H<sub>71</sub>N<sub>24</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1239.6; found: 1239.6

Polyamide **11**: MALDI-TOF MS C<sub>56</sub>H<sub>71</sub>N<sub>24</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup>: 1239.6; found: 1239.5

Polyamide **12**: MALDI-TOF MS  $C_{56}H_{71}N_{24}O_{10}^{+}$  calculated  $[M+H]^{+}$ : 1222.6; found: 1222.7

Polyamide **13**: MALDI-TOF MS  $C_{56}H_{70}N_{23}O_{10}^{+}$  calculated  $[M+H]^{+}$ : 1224.6; found: 1224.6

Polyamide **14**: MALDI-TOF MS  $C_{56}H_{71}N_{24}O_{10}^{+}$  calculated  $[M+H]^{+}$ : 1239.6; found: 1239.7

Polyamide **15**: MALDI-TOF MS  $C_{56}H_{71}N_{24}O_{10}^{+}$  calculated  $[M+H]^{+}$ : 1239.6; found: 1239.4

Polyamide **16**: MALDI-TOF MS  $C_{56}H_{71}N_{24}O_{10}^{+}$  calculated  $[M+H]^{+}$ : 1239.6; found: 1239.5

Polyamide **17**: MALDI-TOF MS  $C_{59}H_{74}N_{23}O_{11}^{+}$  calculated  $[M+H]^{+}$ : 1280.6; found: 1280.6

Polyamide **18**: MALDI-TOF MS  $C_{59}H_{74}N_{23}O_{11}^{+}$  calculated  $[M+H]^{+}$ : 1280.6; found: 1280.7

Polyamide **19**: MALDI-TOF MS  $C_{59}H_{74}N_{23}O_{11}^{+}$  calculated  $[M+H]^{+}$ : 1280.6; found: 1280.6

Polyamide **20**: MALDI-TOF MS  $C_{65}H_{77}N_{22}O_{12}^{+}$  calculated  $[M+H]^{+}$ : 1357.7; found: 1357.6

Polyamide **21**: MALDI-TOF MS  $C_{65}H_{77}N_{22}O_{12}^{+}$  calculated  $[M+H]^{+}$ : 1357.6; found: 1357.7

Polyamide **22**: MALDI-TOF MS  $C_{64}H_{76}N_{23}O_{12}^{+}$  calculated  $[M+H]^{+}$ : 1358.6; found: 1358.6

Polyamide **23**: MALDI-TOF MS  $C_{64}H_{76}N_{23}O_{12}^{+}$  calculated  $[M+H]^{+}$ : 1358.6; found: 1358.6

### Construction of Plasmids pCDMF-1 and pCDMF-2

Oligonucleotides were purchased from Integrated DNA Technologies. The plasmids pCDMF-1 and pCDMF-2 were constructed by annealing the oligonucleotides: 5'-AGCTGCGGCTCGAGACGGCTAACCCATCGAGACGGCTAGCCCATCGAGACGGCTATCCCATCGAGACGGCTACCCCATCGAGAGGATC-3' and 5'-GATCGATCCTCTCGATGGGGTAGCCGTCTCGATGGGATAGCCGTCTCGATGGGCTAGCCGTCTCGATGGGTAGCCGTCTCGAGCCGC-3'; 5'-AGCTGCGAGACGGCTCGAGACGGCTTGAACATCGAGACGGCTCGAGACGGCTTGACCATCGAGACGGCTCGAGACGGCTC-3' and 5'-GATCGAGCCGTCTCGAGCCGTCTCGATGGTCAAGCCGTCTCGAGCCGTCTCGC-3', respectively, followed by ligation into the BamHI/HindIII restriction fragment of pUC19 using T4 DNA ligase.

The plasmid was then transformed into *Escherichia coli* JM109 competent cells. Ampicillin-resistant white colonies were selected from 25 mL Luria–Bertani (LB) agar plates containing 50 mg/mL ampicillin treated with XGAL and isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) solutions and grown overnight at 37 °C. Cells were harvested the following day and purification of the plasmid was performed with a Wizard Plus Midiprep DNA purification kit (Promega). DNA sequencing of the plasmid insert was performed by the sequence analysis facility at the California Institute of Technology.

### DNase I footprinting titrations

Polyamide equilibrations and DNase I footprint titrations were conducted on the 5' end-labeled PCR product of pCDMF-1 and pCDMF-2 according to standard protocols.<sup>34</sup> DNA was incubated with polyamides or water (control) for 12 h at room temperature prior to reaction with DNase I.

### **UV absorption spectrophotometry**

Melting temperature analysis was performed on a Varian Cary 100 spectrophotometer equipped with a thermo-controlled cell holder possessing a cell path length of 1 cm. The buffer for the spectroscopic measurements was chosen to match as closely as possible the conditions of DNase I footprinting experiments. We used 10 mM sodium cacodylate since the temperature dependence of Tris-HCl makes it poorly suited for melting temperature analyses.<sup>37</sup> A degassed aqueous solution of 10 mM sodium cacodylate, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0 was used as analysis buffer. DNA duplexes and hairpin polyamides were mixed in 1:1 stoichiometry to a final concentration of 2  $\mu$ M for each experiment. Prior to analysis, samples were heated to 90 °C and cooled to a starting temperature of 25 °C with a heating rate of 5 °C/min for each ramp. Denaturation profiles were recorded at  $\lambda = 260$  nm from 25 to 90 °C with a heating rate of 0.5 °C/min. The reported melting temperatures were defined as the maximum of the first derivative of the denaturation profile.

### **Molecular modeling**

DNA/polyamide models are based on coordinates derived from NMR structure studies using standard bond length and angles.<sup>8</sup> The molecular graphics images are non-minimized and have been created by introducing ammonium residues to the appropriate position of the turn unit using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081).<sup>49</sup>

### **Measurement of androgen-induced PSA mRNA**

Experiments were performed as described previously.<sup>2</sup>

## Acknowledgments

This work was supported by the National Institutes of Health (GM27681). C.D. is grateful to the Alexander von Humboldt foundation for a postdoctoral research fellowship. M.E.F. is grateful for a pre-doctoral NIH NRSA training grant. D.M.C. is grateful to the Kanel Foundation for a pre-doctoral fellowship.

## References

1. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2004**. *101*, 16768-16773.
2. Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2007**. *104*, 10418-10423.
3. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**. *35*, 363-370.
4. Dervan, P. B. *Bioorg. Med. Chem.* **2001**. *9*, 2215-2235.
5. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**. *13*, 284-299.
6. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science*. **1998**. *282*, 111-115.
7. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C. *Nat. Struct. Biol.* **1998**. *5*, 104-109.
8. Zhang, Q.; Dwyer, T. J.; Tsui, V.; Case, D. A.; Cho, J. H.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **2004**. *126*, 7958-7966.
9. Puckett, J. W.; Muzikar, K. A.; Tietjen, J.; Warren, C. L.; Ansari, A. Z.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**. *129*, 12310-12319.
10. Wurtz, N. R.; Dervan, P. B. *Chem. Biol.* **2000**. *7*, 153-161.
11. Sasaki, S.; Bando, T.; Minoshima, M.; Shimizu, T.; Shinohara, K.; Takaoka, T.; Sugiyama, H. *J. Am. Chem. Soc.* **2006**. *128*, 12162-12168.
12. Tsai, S. M.; Farkas, M. E.; Chou, C. J.; Gottesfeld, J. M.; Dervan, P. B. *Nucleic Acids Res.* **2007**. *35*, 307-316.

13. Minoshima, M.; Bando, T.; Sasaki, S.; Shinohara, K.; Shimizu, T.; Fujimoto, J.; Sugiyama, H. *J. Am. Chem. Soc.* **2007**. *129*, 5384-5390.
14. Poulin-Kerstien, A. T.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**. *125*, 15811-15821.
15. Fechter, E. J.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**. *125*, 8476-8485.
16. Fechter, E. J.; Olenyuk, B.; Dervan, P. B. *Angew. Chem., Int. Ed.* **2004**. *43*, 3591-3594.
17. Fechter, E. J.; Olenyuk, B.; Dervan, P. B. *J. Am. Chem. Soc.* **2005**. *127*, 16685-16691.
18. Rucker, V. C.; Foister, S.; Melander, C.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**. *125*, 1195-1202.
19. Chenoweth, D. M.; Viger, A.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**. *129*, 2216-2217.
20. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B. *Angew. Chem., Int. Ed.* **2007**. *46*, 7956-7959.
21. Schmidt, T. L.; Nandi, C. K.; Rasched, G.; Parui, P. P.; Brutschy, B.; Famulok, M.; Heckel, A. *Angew. Chem., Int. Ed.* **2007**. *46*, 4382-4384.
22. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B. *J. Am. Chem. Soc.* **2008**. *130*, 402-403.
23. Arndt, H. D.; Hauschild, K. E.; Sullivan, D. P.; Lake, K.; Dervan, P. B.; Ansari, A. Z. *J. Am. Chem. Soc.* **2003**. *125*, 13322-13323.
24. Kwonj, Y.; Arndt, H. D.; Qian, M.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M. *J. Am. Chem. Soc.* **2004**. *126*, 15940-15941.
25. Hauschild, K. E.; Metzler, R. E.; Arndt, H. D.; Moretti, R.; Raffaele, M.; Dervan, P. B.; Ansari, A. Z. *Proc. Natl. Acad. Sci. U.S.A.* **2005**. *102*, 5008-5013.
26. Stafford, R. L.; Arndt, H. D.; Brezinski, M. L.; Ansari, A. Z.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**. *129*, 2660-2668.
27. Stafford, R. L.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**. *129*, 14026-14033.
28. Xiao, X. S.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T. *Angew. Chem., Int. Ed.* **2007**. *46*, 2865-2868.
29. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**. *116*, 7983-7988.

30. Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**. *120*, 1382-1391.
31. Zhang, W.; Minoshima, M.; Sugiyama, H. *J. Am. Chem. Soc.* **2006**. *128*, 14905-14912.
32. Farkas, M. E.; Tsai, S. M.; Dervan, P. B. *Bioorg. Med. Chem.* **2007**. *15*, 6927-6936.
33. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B. Z.; Puckett, J. W.; Wang, C. C. C.; Dervan, P. B. *Tetrahedron*. **2007**. *63*, 6146-6151.
34. Trauger, J. W.; Dervan, P. B. *Meth. Enzymol.* **2001**. *340*, 450-466.
35. Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Method Enzymol.* **1986**. *130*, 132-181.
36. Senear, D. F.; Dalmawiszhausz, D. D.; Brenowitz, M. *Electrophoresis*. **1993**. *14*, 704-712.
37. Pilch, D. S.; Poklar, N.; Gelfand, C. A.; Law, S. M.; Breslauer, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1996**. *93*, 8306-8311.
38. Pilch, D. S.; Poklar, N.; Baird, E. E.; Dervan, P. B.; Breslauer, K. J. *Biochemistry*. **1999**. *38*, 2143-2151.
39. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**. *120*, 6219-6226.
40. Floreancig, P. E.; Swalley, S. E.; Trauger, J. W.; Dervan, P. B. *J. Am. Chem. Soc.* **2000**. *122*, 6342-6350.
41. Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**. *121*, 1113-1120.
42. Weyermann, P.; Dervan, P. B. *J. Am. Chem. Soc.* **2002**. *124*, 6872-6878.
43. Herman, D. M.; Baird, E. E.; Dervan, P. B. *Chem.-Eur. J.* **1999**. *5*, 975-983.
44. Wang, C. C. C.; Dervan, P. B. *J. Am. Chem. Soc.* **2001**. *123*, 8657-8661.
45. Edayathumangalam, R. S.; Weyermann, P.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K. *Proc. Natl. Acad. Sci. U. S. A.* **2004**. *101*, 6864-6869.
46. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003**. *100*, 12063-12068.



47. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004.** 32, 2802-2818.
48. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996.** 118, 6141-6146.
49. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. **2004.** 25, 1605-1612.

## Chapter IIIB

### DNA sequence selectivity of hairpin polyamide turn units

*The text of this chapter was taken in part from a manuscript co-authored with Benjamin C. Li, Christian Dose, and Peter B. Dervan (California Institute of Technology)*

(Farkas, M.E.; Li, B.C.; Dose, C.; Dervan, P.B. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6859-6866)

**Abstract**

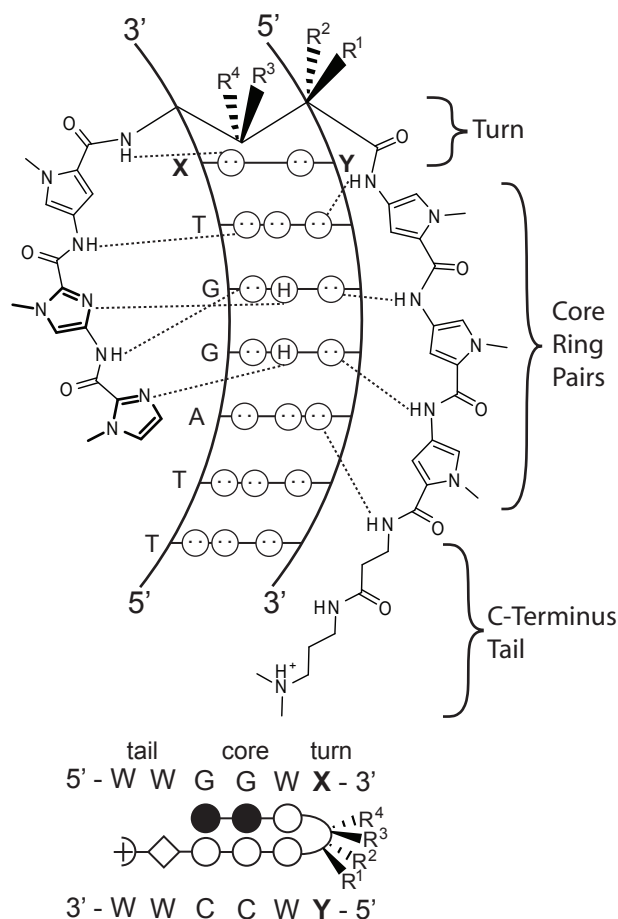
A class of hairpin polyamides linked by 3,4-diaminobutyric acid, resulting in a  $\beta$ -amine residue at the turn unit, showed improved binding affinities relative to their  $\alpha$ -amino- $\gamma$ -turn analogs for particular sequences. We incorporated  $\beta$ -amino- $\gamma$ -turns in six-ring polyamides and determined whether there are any sequence preferences under the turn unit by quantitative footprinting titrations. Although there was an energetic penalty for G•C and C•G base pairs, we found little preference for T•A over A•T at the  $\beta$ -amino- $\gamma$ -turn position. Fluorine and hydroxyl substituted  $\alpha$ -amino- $\gamma$ -turns were synthesized for comparison. Their binding affinities and specificities in the context of six-ring polyamides demonstrated overall diminished affinity and no additional specificity at the turn position. We anticipate that this study will be a baseline for further investigation of the turn subunit as a recognition element for the DNA minor groove.

## Introduction

Hairpin pyrrole-imidazole (Py/Im) polyamides are a class of programmable synthetic ligands able to bind a broad repertoire of DNA sequences with affinities and specificities comparable to those of DNA-binding proteins.<sup>1,2</sup> They have been shown to localize to the nuclei of living cells<sup>3,4</sup> and regulate endogenous gene expression by interfering with transcription factor/DNA interfaces.<sup>5-10</sup> Discrimination of the four Watson-Crick base pairs is dependent upon Py/Im ring pairings in the minor groove. Pairing rules have been established whereby *N*-methyl imidazole/*N*-methylpyrrole (Im/Py) pairs target G•C, the reverse (Py/Im) target C•G, and Py/Py pairs target A•T and T•A.<sup>11-13</sup>

The turn unit in the hairpin is a recognition element, favoring T•A / A•T over G•C / C•G.<sup>14</sup> There is an energetic penalty for unfavorable steric interaction with the exocyclic amine present at the edge of the G•C base pair (Figure IIIB.1). The question arises whether discrimination between T•A and A•T with the turn unit can be achieved. Previous efforts toward the improvement of hairpin binding affinity have involved modifications of the turn unit. Early studies showed the optimal length of the turn element to be three methylene units, resulting in the use of  $\gamma$ -aminobutyric acid ( $\gamma$ -turn).<sup>15</sup> Modification of the  $\alpha$ -position of the parent  $\gamma$ -turn with (*R*)-2,4-diaminobutyric acid ( $\alpha$ -amino- $\gamma$ -turn) results in an approximately 15-fold increase in DNA-binding affinity.<sup>16</sup> In contrast, hairpins containing the opposite enantiomer, (*S*)- $\alpha$ -amino- $\gamma$ -turn, bind DNA with diminished affinities likely due to a steric clash of the amine with the wall of the minor groove. Polyamides containing  $\alpha$ -hydroxy- $\gamma$ -turns<sup>17</sup> and  $\alpha$ -diaminobutyric acid<sup>18</sup> turns have been reported to impart additional elements of specificity, but with the cost of diminished binding affinities.

In a formal sense, a DNA minor groove binding hairpin Py/Im polyamide is an early example of a class of oligomers encoded by the order of monomer units that fold to a desired shape with a specific function, referred to as “foldamers.”<sup>19-21</sup> The ring order of Py/Im polyamides codes in a programmable manner for a specific, contiguous sequence of Watson-Crick base pairs.<sup>22</sup> The turn unit is both a shape element as well as a DNA



**Figure IIB.1.** Schematic diagram of six-ring hairpin polyamide (ImImPy-turn-PyPyPy) targeting the DNA sequence 5'-WWGGWW-3'. Dashed lines indicate hydrogen bonds between the polyamide and DNA base pairs.

recognition element, allowing the molecule to fold in a U-conformation. The turn unit deserves more attention and this paper represents an effort to create a baseline for the field.

Recently, we introduced a new class of hairpin polyamides linked by 3,4-diaminobutyric acid, resulting in a  $\beta$ -amino- $\gamma$ -turn.<sup>23</sup> These molecules showed improved binding affinities relative to their  $\alpha$ -amino- $\gamma$ -turn analogs for A/T-rich sequences. Additionally, polyamides containing the  $\beta$ -amino- $\gamma$ -turn were found to have improved tolerance for synthetic modification at the amine position presumably due to their more

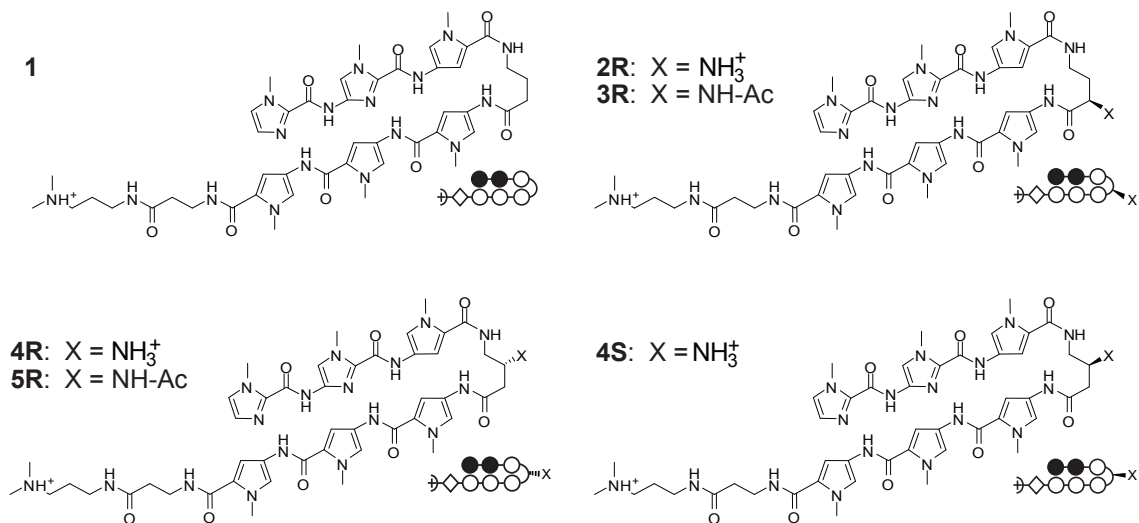
central location on the floor of the minor groove. Due to limitations of quantitative DNase I footprinting titrations<sup>24-26</sup> for compounds where  $K_a \geq 10^{10} \text{ M}^{-1}$ , relative binding affinities for high affinity molecules are compared by using their thermal stabilization of DNA duplexes. It has been shown previously that increases in melting temperatures ( $\Delta T_m$ ) of DNA duplexes bound by hairpin polyamides correlate with DNA-binding affinity.<sup>27,28</sup>

We report herein a comparison of the sequence specificities of hairpin Py/Im polyamides containing the  $\alpha$ -amino- $\gamma$ -turn and the  $\beta$ -amino- $\gamma$ -turn. Additionally, we have synthesized both hydroxyl and fluoro-substituted  $\alpha$ -amino- $\gamma$ -turns and determined their affinities and specificities in polyamides with analogous core ring pairs. By employing six-ring polyamides, which have lower binding affinities compared to eight-ring polyamides,<sup>29</sup> we are able to determine reliable equilibrium association constants ( $K_a$ ) via quantitative DNase I footprinting titrations, and compare them with DNA duplex thermal stabilizations.

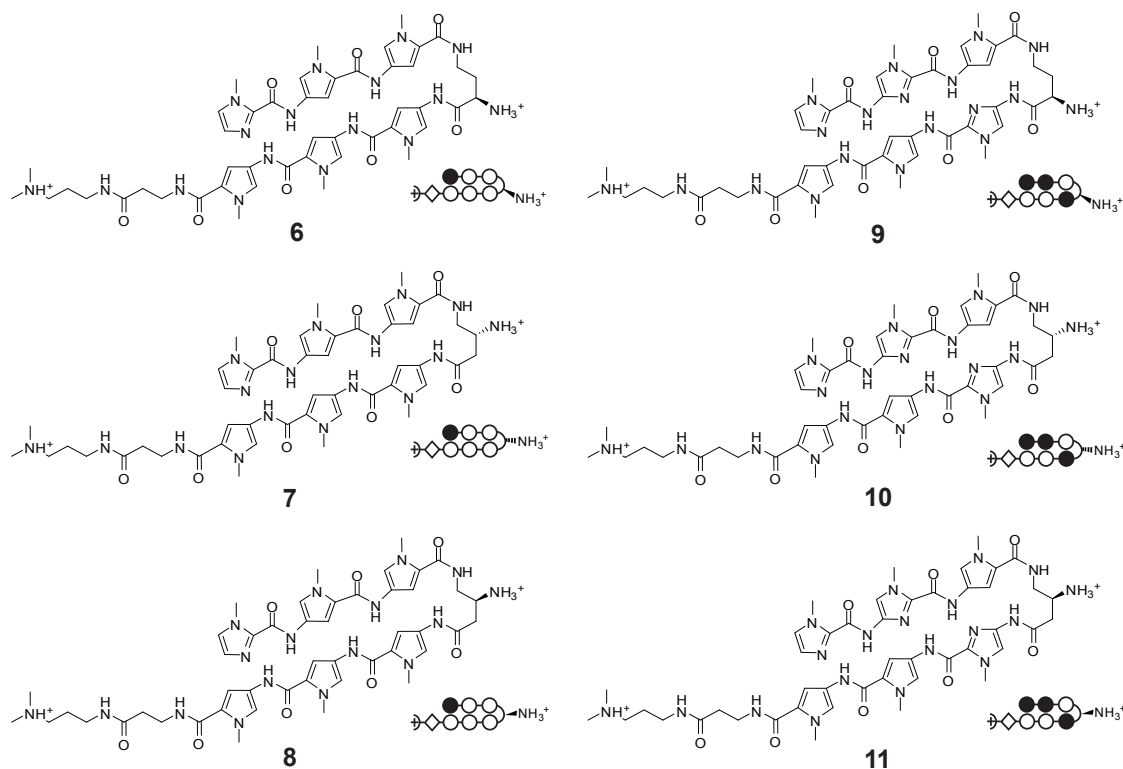
## Results and Discussion

### Polyamide synthesis

Six-ring hairpin polyamides (ImImPy-turn-PyPyPy) targeting the DNA sequence 5'-WWGGWW-3' were synthesized by solid phase methods on Pam resin (Figure IIIB.2).<sup>23,30</sup> In addition to the parent molecule containing an unsubstituted  $\gamma$ -aminobutyric acid hairpin (**1**), oligomers containing an amine moiety in the  $\alpha$  and  $\beta$  turn positions (**2R**, **4R**, **4S**) and two polyamides with acetylated amines (**3R**, **5R**) were also synthesized. The acetylated, or capped, molecules were used to determine tolerance for modifications at the turn. Acetylated turn units have been shown to improve nuclear uptake of polyamides.<sup>4</sup> In order to expand the repertoire of molecules studied, two other series of polyamides, **6-11** (Figure IIIB.3), targeting the DNA sequences 5'-WWGWWW-3' and 5'-WWGGCW-3' were also synthesized. For these series, only molecules containing amines in the  $\alpha$  and  $\beta$  turn positions were made.



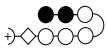
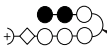
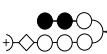

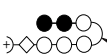
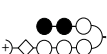
**Figure IIIB.2.** Chemical and ball-and-stick structures of polyamides targeting 5'-WWGGWW-3'. Ball and stick symbols are defined as follows: pyrrole is denoted by an open circle, imidazole is denoted by a filled circle, and  $\beta$ -alanine is denoted by a diamond shape.



**Figure IIIB.3.** Chemical and ball-and-stick structures of polyamides **6-11**, targeting the DNA sequences 5'-WWGWWW-3' and 5'-WWGGCW-3'. Ball and stick symbols are defined as follows: pyrrole is denoted by an open circle, imidazole is denoted by a filled circle, and  $\beta$ -alanine is denoted by a diamond shape.


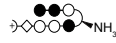

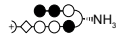
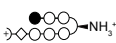

### Thermal stabilization of DNA duplexes

Spectroscopic analyses of the polyamides targeting 5'-WWGGWW-3' were performed on the 11mer DNA duplex shown in Table IIIB.1.<sup>23</sup> All hairpins analyzed provided an increase in melting temperature, confirming the formation of DNA/polyamide complexes. The  $\Delta T_m$  values obtained for polyamides containing a free amine were within error of each other. However, acetylated polyamide **3R** ( $\alpha$ -amino turn) showed a greater decrease in affinity than **5R** ( $\beta$ -amino- $\gamma$ -turn). As had been demonstrated with eight-ring hairpin molecules,<sup>23</sup> improvements in binding affinities for  $\beta$ - over  $\alpha$ -amino- $\gamma$ -turn six-ring polyamides are more pronounced with decreasing imidazole content (Table IIIB.2). Polyamides **6-9**, which contain only a single imidazole ring, were analyzed on two different duplexes: one where the binding site is flanked by C•G, and another by T•A, for the purpose of reducing binding affinity in subsequent footprinting analyses.

DNA = 5'-CTAT <b>TGGTA</b> GAC-3'			
Polyamides	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$	
—	45.1 ( $\pm 0.1$ )	—	
<b>1</b> 	54.3 ( $\pm 0.2$ )	9.2	
<b>2</b> 	59.8 ( $\pm 0.3$ )	14.7	
<b>3</b> 	60.0 ( $\pm 0.2$ )	14.9	
<b>4</b> 	59.8 ( $\pm 0.2$ )	14.7	
<b>2-Ac</b> 	54.9 ( $\pm 0.1$ )	10.1	
<b>3-Ac</b> 	57.0 ( $\pm 0.2$ )	12.2	

**Table IIIB.1.** Melting temperatures of DNA/polyamide complexes targeting 5'-WWGGWW-3'. All values are derived from at least three melting temperature experiments, with standard deviations indicated in parentheses.  $\Delta T_m$  values are given as  $T_m^{(\text{DNA/polyamide})} - T_m^{(\text{DNA})}$ .



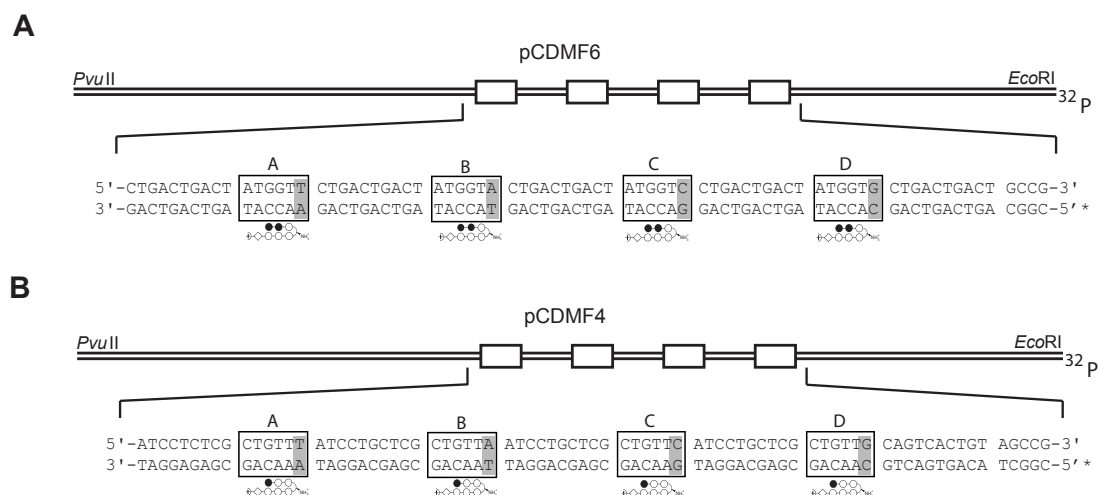
DNA = 5'-CTC <b>TGTTA</b> CCA-3'			DNA = 5'-CTA <b>TGTTA</b> CCA-3'			DNA = 5'-CTC <b>TGGCA</b> CAC-3'		
Polyamides	T <sub>m</sub> / °C	ΔT <sub>m</sub> / °C	T <sub>m</sub> / °C	ΔT <sub>m</sub> / °C		Polyamides	T <sub>m</sub> / °C	ΔT <sub>m</sub> / °C
—	47.3 (±0.3)	—	44.0 (±0.1)	—		—	56.0 (±0.2)	—
<b>6</b> 	58.2 (±0.4)	10.9	59.9 (±0.3)	15.9		<b>9</b> 	64.3 (±0.3)	8.3
<b>7</b> 	60.3 (±0.3)	13.0	62.0 (±0.1)	18.0		<b>10</b> 	62.7 (±0.3)	6.7
<b>8</b> 	58.9 (±0.2)	11.6	61.0 (±0.2)	17.0		<b>11</b> 	62.0 (±0.3)	6.0

**Table IIIB.2.** Melting temperatures of DNA/polyamide complexes. All values are derived from at least three melting temperature experiments, with standard deviations indicated in parentheses. ΔT<sub>m</sub> values are given as T<sub>m</sub><sup>(DNA/polyamide)</sup> - T<sub>m</sub><sup>(DNA)</sup>.

### DNA binding affinity and sequence selectivity

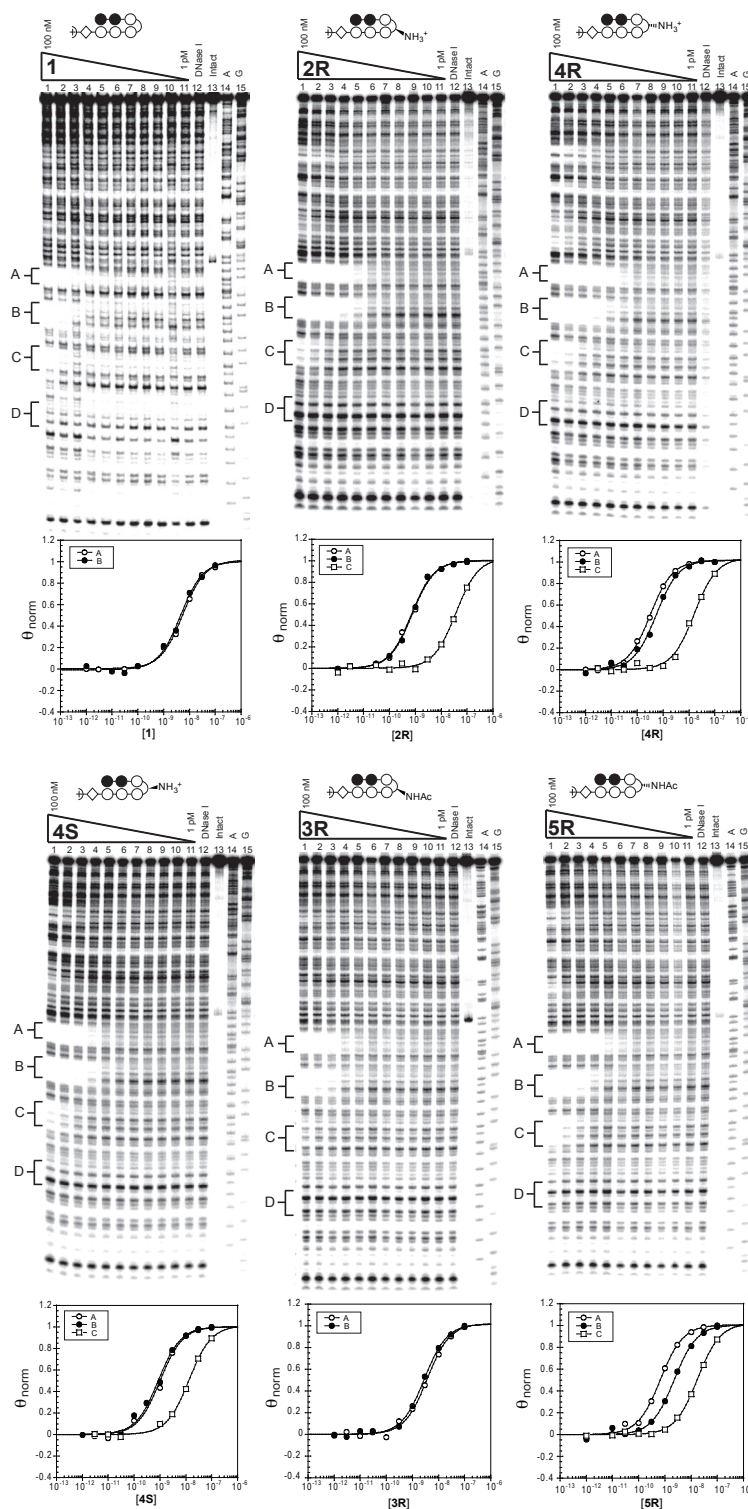
The plasmid pCDMF6 was prepared to characterize polyamides targeting the sequence 5'-WWGGWW-3' (Figure IIIB.4A), while pCDMF4 was prepared to characterize polyamides targeting 5'-WWGWWW-3' (Figure IIIB.4B). The designed inserts each contain four binding sites, varying the nucleotide base pair present under the turn unit. In order to avoid the compression of binding affinities ( $K_a$ ) for polyamides **6-8**, pCDMF4 was designed to have C•G base pairs under the polyamide dimethylpropylamine tail unit, diminishing binding. Quantitative DNase I footprinting titrations were performed with the polyamides in order to measure their binding site affinities and specificities, as previously described.<sup>25</sup>

As expected, the parent hairpin containing the γ-turn retained the lowest binding affinity, while experiments for **2R**, **4R**, and **4S** (Figure IIIB.5, Table IIIB.3) corroborated the similar ΔT<sub>m</sub> values obtained for duplex stabilization. None of these molecules bound the G•C base pair, and binding to C•G was greatly diminished. Polyamide **4R** revealed only a ~2-fold specificity for T•A over A•T (Table IIIB.4). Analysis of polyamides **3R** and **5R** revealed a greater decrease in binding affinities for the α-amino-γ-turn molecule



**Figure IIIB.4.** Illustration of the EcoRI/PvuII restriction fragment derived from plasmid (A) pCDFM6 and (B) pCDFM4, used to characterize polyamides. The designed polyamide binding sites are indicated by boxes; single base-pair mismatches are indicated by shaded regions.

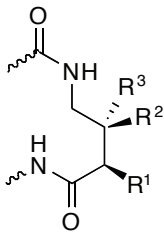
than the  $\beta$ . Between **2R** and **3R** there is a five-fold decrease over T•A, and an eight-fold decrease over A•T. For  $\beta$ -amino- $\gamma$ -turn polyamides **4R** and **5R**, there are 1.7 and 3.2-fold decreases for T•A and A•T, respectively. **5R** shows ~3-fold preference for binding T•A versus A•T (Table IIIB.4). Studies of polyamides targeting 5'-WWGWWW-3' revealed similar trends (Figure IIIB.6, Table IIIB.5 and 6).



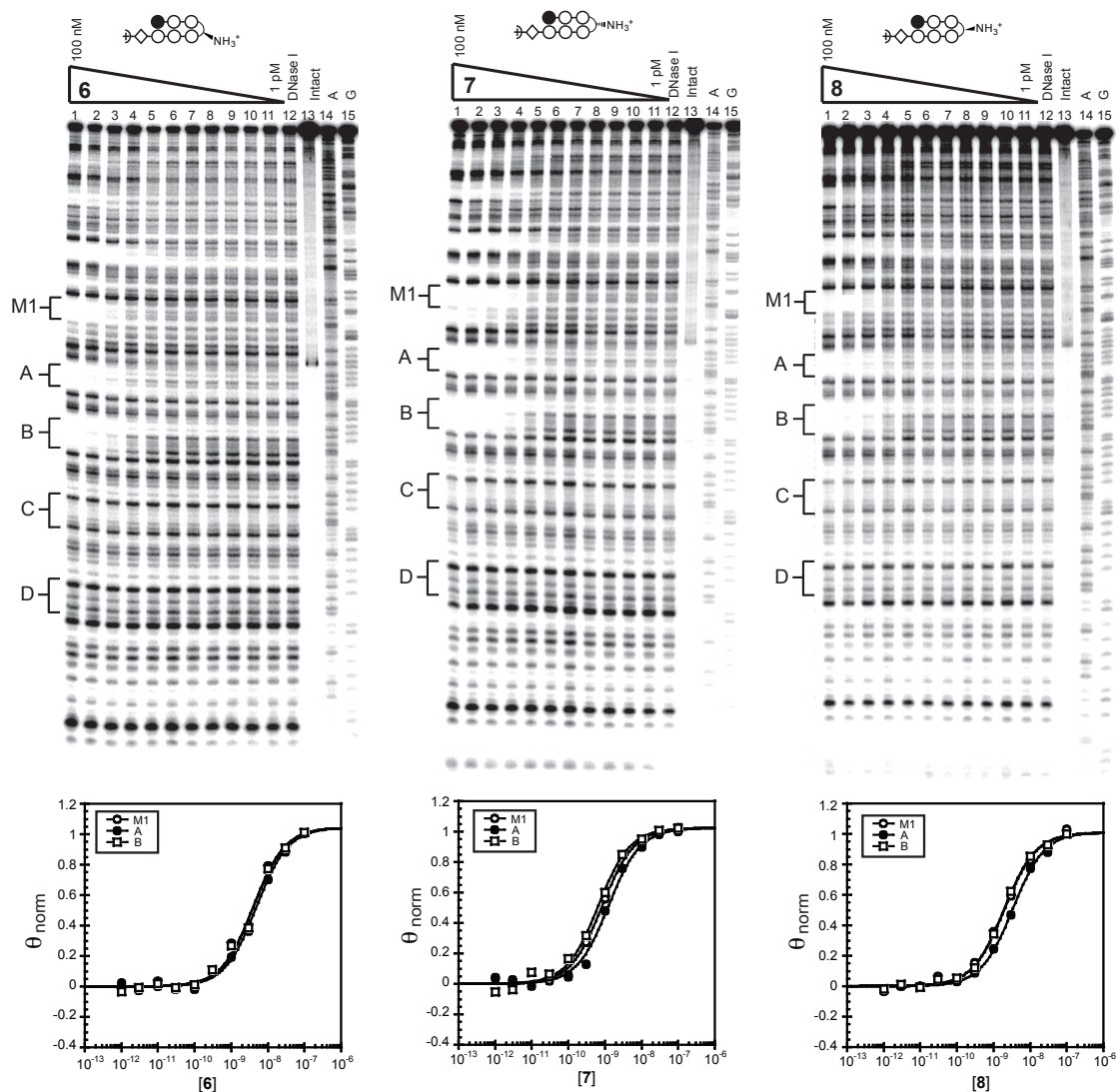
**Figure IIIB.5.** Quantitative DNase I footprinting titration experiments for polyamides **1**, **2R**, and **4R** (top, left to right) and **4S**, **3R**, and **5R** (bottom, left to right) on the 5' end labeled PCR product of plasmid pCD-MF6: lanes 1-11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Respective isotherms are shown below.

	5'-ATGGTT-3'	5'-ATGGTA-3'	5'-ATGGTC-3'	5'-ATGGTG-3'
<b>1</b>	2.1 ( $\pm$ 0.6) $\times 10^8$	2.0 ( $\pm$ 0.5) $\times 10^8$	< $10^7$	< $10^7$
<b>2R</b>	1.6 ( $\pm$ 0.7) $\times 10^9$	1.7 ( $\pm$ 0.9) $\times 10^9$	3.9 ( $\pm$ 2.0) $\times 10^7$	< $10^7$
<b>4R</b>	3.3 ( $\pm$ 0.1) $\times 10^9$	1.8 ( $\pm$ 0.2) $\times 10^9$	5.0 ( $\pm$ 1.1) $\times 10^7$	< $10^7$
<b>4S</b>	1.1 ( $\pm$ 0.1) $\times 10^9$	1.3 ( $\pm$ 0.2) $\times 10^9$	6.0 ( $\pm$ 0.4) $\times 10^7$	< $10^7$
<b>3R</b>	3.2 ( $\pm$ 0.4) $\times 10^8$	2.1 ( $\pm$ 0.3) $\times 10^8$	< $10^7$	< $10^7$
<b>5R</b>	2.0 ( $\pm$ 0.4) $\times 10^9$	5.2 ( $\pm$ 1.6) $\times 10^8$	8.9 ( $\pm$ 2.8) $\times 10^7$	< $10^7$

**Table IIIB.3.** Binding affinities for polyamides targeting 5'-WWGGWW-3'. Equilibrium association constants ( $M^{-1}$ ) are reported as mean values from three DNase I footprinting titration experiments. Standard deviations are shown in parentheses.

	R <sup>1</sup> , R <sup>2</sup> , R <sup>3</sup>	T • A	A • T	C • G	G • C
	H, H, H	1.0	1.0	< 0.01	< 0.01
	NH <sub>3</sub> <sup>+</sup> , H, H	1.0	0.9	0.02	< 0.01
	H, H, NH <sub>3</sub> <sup>+</sup>	1.0	0.5	0.02	< 0.01
	H, NH <sub>3</sub> <sup>+</sup> , H	1.0	1.2	0.05	< 0.01
	NH-Ac, H, H	1.0	0.7	< 0.01	< 0.01
	H, H, NH-Ac	1.0	0.3	0.04	< 0.01

**Table IIIB.4.** Relative binding affinities for polyamides. Relative binding affinities are reported as ratios of binding affinities ( $K_a$ ) as determined by DNase I footprinting titration experiments for polyamides targeting 5'-WWGGWW-3'. Affinities less than  $10^7$  cannot be determined.



**Figure IIIB.6.** Quantitative DNase I footprinting titration experiments on plasmid pCDMF4 for polyamides 6, 7, and 8. Lanes 1-11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Respective binding isotherms are shown below.

	5'-CTGTTT-3'	5'-CTGTTA-3'	5'-ATGATT-3'	5'-CTGTTC-3'	5'-CTGTTG-3'
<b>6</b>	$2.0 (\pm 0.1) \times 10^8$	$2.8 (\pm 0.3) \times 10^8$	$2.8 (\pm 0.2) \times 10^8$	$< 10^7$	$< 10^7$
<b>7</b>	$8.2 (\pm 0.2) \times 10^8$	$2.2 (\pm 1.5) \times 10^9$	$1.3 (\pm 0.3) \times 10^9$	$< 10^7$	$< 10^7$
<b>8</b>	$3.2 (\pm 0.2) \times 10^8$	$5.3 (\pm 0.5) \times 10^8$	$5.0 (\pm 0.4) \times 10^8$	$< 10^7$	$< 10^7$

**Table IIIB.5.** Binding affinities for polyamides targeting 5'-WWGWW-3'. Equilibrium association constants ( $M^{-1}$ ) are reported as mean values from three DNase I footprinting titration experiments. Standard deviations are shown in parentheses. Affinities less than  $10^7$  cannot be determined.

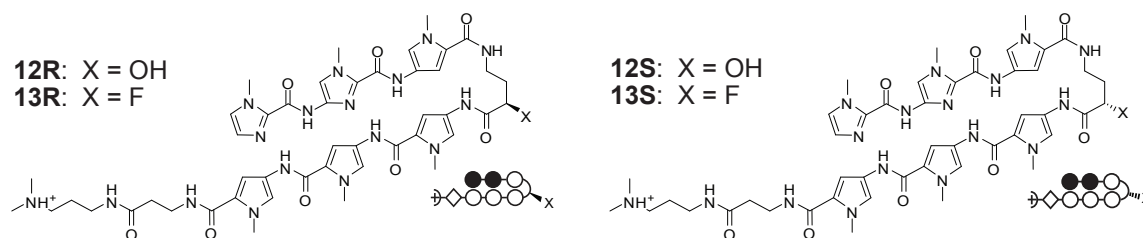
	$R^1, R^2, R^3$	$\begin{matrix} T \\ \cdot \\ A \end{matrix}$	$\begin{matrix} A \\ \cdot \\ T \end{matrix}$	$\begin{matrix} C \\ \cdot \\ G \end{matrix}$	$\begin{matrix} G \\ \cdot \\ C \end{matrix}$
	$NH_3^+, H, H$	<b>1.0</b>	<b>1.4</b>	<b>&lt; 0.01</b>	<b>&lt; 0.01</b>
	$H, H, NH_3^+$	<b>1.0</b>	<b>2.7</b>	<b>&lt; 0.01</b>	<b>&lt; 0.01</b>
	$H, NH_3^+, H$	<b>1.0</b>	<b>1.6</b>	<b>&lt; 0.01</b>	<b>&lt; 0.01</b>

**Table IIIB.6.** Relative binding affinities for polyamides. Relative binding affinities are reported as ratios of binding affinities ( $K_a$ ) as determined by DNase I footprinting titration experiments for polyamides targeting 5'-WWGWW-3'.

### Fluoro and hydroxyl substituted turn units

$\alpha$ -Fluoro and hydroxyl substituted hairpin turns were synthesized and incorporated in six-ring hairpin polyamides targeting the DNA sequence 5'-WWGGWW-3' (Figure IIIB.7). Both (*R*) and (*S*) enantiomers of each turn subunit were examined. Polyamides were synthesized on Pam resin using standard solid phase methods.

Thermal stabilization analysis of the polyamides on 11mer DNA duplexes revealed that fluoro and hydroxyl substituted hairpins resulted in lower stabilizations than the corresponding amine-substituted and acetylated polyamides (Table IIIB.7). The presence of T•A and A•T base pairs under the hairpin turn subunit were probed. DNase I footprinting titrations on the plasmid pCDMF6 (Figure IIIB.8, Table IIIB.8) showed that



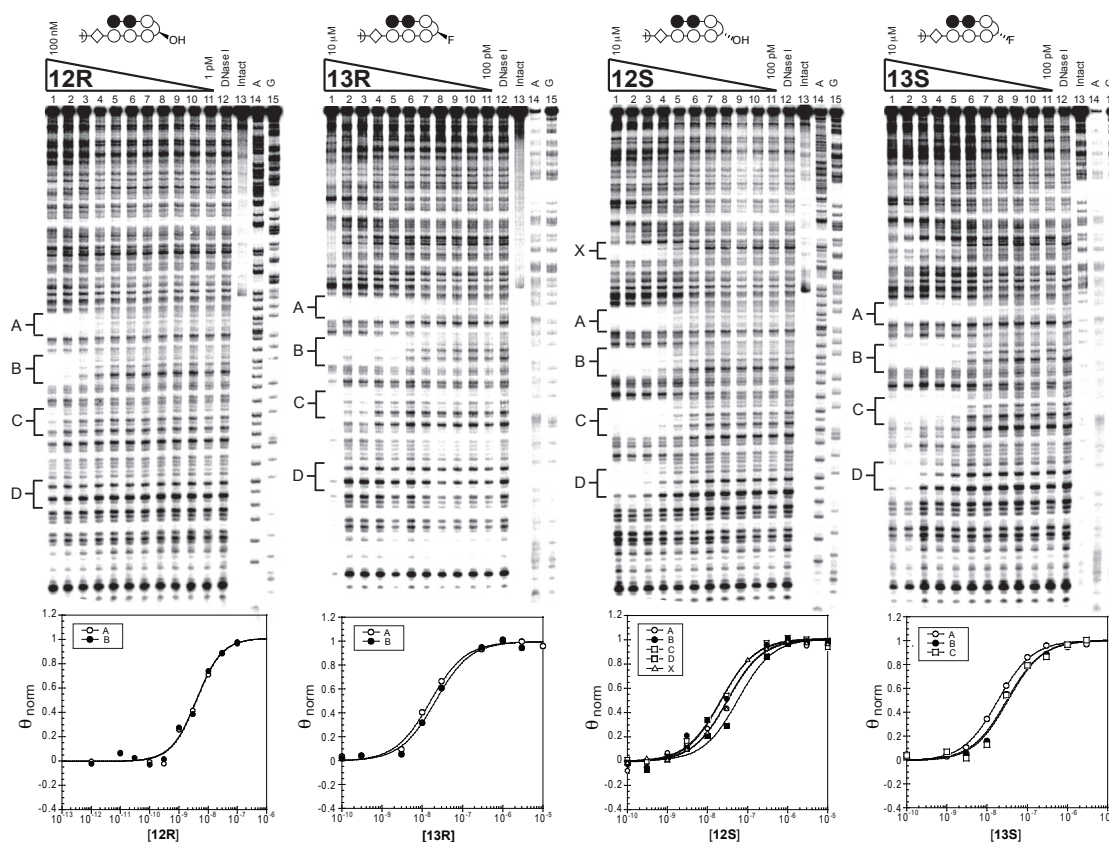
**Figure IIIB.7.** Chemical and ball-and-stick structures of polyamides containing fluoro and hydroxyl substituted hairpin turns. Ball and stick symbols are defined as follows: pyrrole is denoted by an open circle, imidazole is denoted by a filled circle, and  $\beta$ -alanine is denoted by a diamond shape.

Polyamide	DNA = 5'-CTAT <b>TGGTA</b> GAC-3'		DNA = 5'-CTAT <b>TGGTA</b> GAC-3'	
	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$
—	45.1 ( $\pm 0.1$ )	—	46.7 ( $\pm 0.4$ )	—
<b>5</b>	55.1 ( $\pm 0.2$ )	10.0	57.4 ( $\pm 0.4$ )	10.7
<b>6</b>	53.4 ( $\pm 0.2$ )	8.3	55.1 ( $\pm 0.1$ )	8.4
<b>7</b>	50.9 ( $\pm 0.2$ )	5.8	52.3 ( $\pm 0.2$ )	5.6
<b>8</b>	53.0 ( $\pm 0.2$ )	7.9	55.4 ( $\pm 0.2$ )	8.7

**Table IIIB.7.** Melting temperatures of DNA/polyamide complexes. All values are derived from at least three melting temperature experiments, with standard deviations indicated in parentheses.  $\Delta T_m$  values are given as  $T_m^{(\text{DNA/polyamide})} - T_m^{(\text{DNA})}$ .

both enantiomers of the hydroxyl and fluoro-hairpin turns resulted in decreased polyamide binding affinities relative to their amine substituted counterparts. Additionally, none of these subunits resulted in increased elements of specificity at the turn position of the molecule (Table IIIB.9).



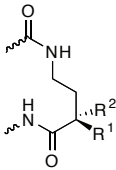


**Figure IIB.8** Quantitative DNase I footprinting titration experiments for polyamides **12R**, **13R**, **12S**, and **13S** (left to right) on the 5' end labeled PCR product of plasmid pCDMF6. For **12R**: lanes 1-11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Polyamides **13R**, **12S**, and **13S**: lanes 1-11, 10  $\mu$ M, 3  $\mu$ M, 1  $\mu$ M, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, and 100 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Respective isotherms are shown below.



	5'-ATGGTT-3'	5'-ATGGTA-3'	5'-ATGGTC-3'	5'-ATGGTG-3'
<b>12R</b>	$2.1 (\pm 0.4) \times 10^8$	$2.0 (\pm 0.5) \times 10^8$	$< 10^7$	$< 10^7$
<b>13R</b>	$5.3 (\pm 1.1) \times 10^7$	$4.5 (\pm 0.6) \times 10^7$	$< 10^7$	$< 10^7$
<b>12S</b>	$2.6 (\pm 0.6) \times 10^7$	$3.4 (\pm 1.2) \times 10^7$	$3.6 (\pm 1.1) \times 10^7$	$2.0 (\pm 0.7) \times 10^7$
<b>13S</b>	$6.1 (\pm 2.3) \times 10^7$	$3.9 (\pm 1.3) \times 10^7$	$3.5 (\pm 1.2) \times 10^7$	$< 10^7$

**Table IIIB.8.** Binding affinities for polyamides containing fluoro- and hydroxyl-substituted turn units. Equilibrium association constants ( $M^{-1}$ ) are reported as mean values from three DNase I footprinting titration experiments. Standard deviations are shown in parentheses. Affinities less than  $10^7$  cannot be determined.

	R <sup>1</sup> ,R <sup>2</sup>	T • A	A • T	C • G	G • C
	OH, H	1.0	1.0	< 0.01	< 0.01
	F, H	1.0	0.8	< 0.01	< 0.01
	H, OH	1.0	1.3	1.4	0.8
	H, F	1.0	0.6	0.6	< 0.01

**Table IIIB.9.** Relative binding affinities for polyamides containing fluoro- and hydroxyl-substituted turn units. Relative binding affinities are reported as ratios of binding affinities ( $K_a$ ) as determined by DNase I footprinting titration experiments for polyamides.

## Conclusion

By utilizing six-ring hairpin polyamides, we were able to combine DNase I footprinting titration and duplex stabilization analyses in order to fully characterize the binding preferences of various hairpin turn subunits. Although the hairpin turns investigated herein show modest DNA binding specificities, we anticipate that further study will yield moieties enabling discrimination amongst all four Watson-Crick base pairs and add an additional recognition element for DNA recognition.

## Materials and Methods

### Polyamide synthesis

Polyamides were synthesized with Boc- $\beta$ -Ala-PAM resin (Peptides International) according to published synthesis protocols.<sup>23,30</sup> Boc-4-Abu-OH was purchased from Novabiochem, (*R*)-2,4-Fmoc-Dbu(Boc)-OH was purchased from Bachem. (*R*)-3,4-CBZ-Dbu(Boc)-OH and (*S*)-3,4-CBZ-Dbu(Boc)-OH were purchased from Senn Chemicals AG. Generation of the hydroxyl and fluoro hairpin turns is described below. All DNA oligomers were purchased from Integrated DNA Technologies. Polyamides were purified via preparative HPLC on an Agilent Technologies 1200 Series system using a Phenomenex Gemini column. Matrix-assisted, LASER desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on an Applied Biosystems Voyager DR Pro spectrometer.

Polyamide **1**: MALDI-TOF MS  $C_{46}H_{60}N_{17}O_8^+$  calculated  $[M+H]^+$ : 979.1; found: 978.8

Polyamide **2R**: MALDI-TOF MS  $C_{46}H_{62}N_{18}O_8^+$  calculated  $[M+H]^+$ : 994.1; found: 993.9

Polyamide **3R**: MALDI-TOF MS  $C_{48}H_{63}N_{18}O_9^+$  calculated  $[M+H]^+$ : 1036.1; found: 1036.0

Polyamide **4R**: MALDI-TOF MS  $C_{46}H_{62}N_{18}O_8^+$  calculated  $[M+H]^+$ : 994.1; found: 994.0

Polyamide **5R**: MALDI-TOF MS  $C_{48}H_{63}N_{18}O_9^+$  calculated  $[M+H]^+$ : 1036.1; found: 1035.9

Polyamide **4S**: MALDI-TOF MS  $C_{46}H_{62}N_{18}O_8^+$  calculated  $[M+H]^+$ : 994.1; found: 993.9

Polyamide **6**: MALDI-TOF MS  $C_{47}H_{62}N_{17}O_8^+$  calculated  $[M+H]^+$ : 992.5; found: 992.2

Polyamide **7**: MALDI-TOF MS  $C_{47}H_{62}N_{17}O_8^+$  calculated  $[M+H]^+$ : 992.5; found: 992.6

Polyamide **8**: MALDI-TOF MS  $C_{47}H_{62}N_{17}O_8^+$  calculated  $[M+H]^+$ : 992.5; found: 992.3

Polyamide **9**: MALDI-TOF MS  $C_{45}H_{60}N_{19}O_8^+$  calculated  $[M+H]^+$ : 995.5; found: 995.0

Polyamide **10**: MALDI-TOF MS  $C_{45}H_{60}N_{19}O_8^+$  calculated  $[M+H]^+$ : 995.5; found: 995.2

Polyamide **11**: MALDI-TOF MS  $C_{45}H_{60}N_{19}O_8^+$  calculated  $[M+H]^+$ : 995.5; found: 995.7

Polyamide **12R**: MALDI-TOF MS  $C_{46}H_{60}N_{17}O_9^+$  calculated  $[M+H]^+$ : 995.1; found: 995.0

Polyamide **13R**: MALDI-TOF MS  $C_{46}H_{59}FN_{17}O_8^+$  calculated  $[M+H]^+$ : 996.5; found: 996.7

Polyamide **12S**: MALDI-TOF MS  $C_{46}H_{60}N_{17}O_9^+$  calculated  $[M+H]^+$ : 995.1; found: 995.1

Polyamide **13S**: MALDI-TOF MS  $C_{46}H_{59}FN_{17}O_8^+$  calculated  $[M+H]^+$ : 996.5; found: 996.6

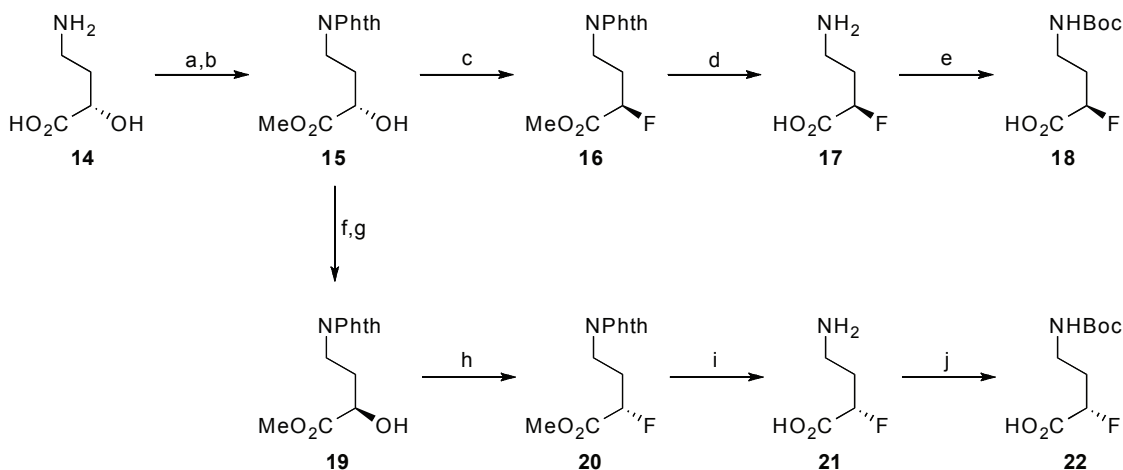
### Polyamide assays

Melting temperature analysis was performed as previously described<sup>22</sup> on a Varian Cary 100 spectrophotometer equipped with a thermo-controlled cell holder, using 10 mM sodium cacodylate buffer.

Plasmids were constructed as previously described<sup>18</sup> by annealing the following oligonucleotide pairs: 5'-AGCTGCTGACTGACTATGGTTCTGACTGACTATGGTACTGACTGACTATGGTCCTGACTGACTATGGTGCTGACTGACTGCCGC-3' and 5'-AGCTGCGGCAGTCAGTCAGCACCATAGTCAGTCAGGACCATAGTCAGTCAGTACCATAGTCAGTCAGAACCATAGTCAGTCAGC-3' for pCDMF6; 5'-AGCTGATCCTCTCGCTGTTTATCCTGCTCGCTGTTAATCCTGCTCGCTGTTTCATCCTGCTCGCTGTTGCAGTCACTGTAGCCGC-3' and 5'-AGCTGCGGCTACAGTGACTGCAACAGCGAGCAGGATGAACAGCGAGCAGGATTAA-CAGCGAGCAGGATAAACAGCGAGAGGATC-3' for pCDMF4. 5'-labeling of primers with <sup>32</sup>P and subsequent DNase I footprinting titrations were performed as previously described.<sup>18,26</sup>

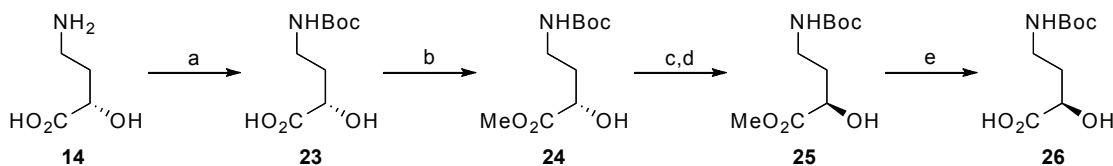
## Synthesis of fluoro and hydroxyl turn precursors

Synthetic schemes for the synthesis of fluorine and hydroxyl substituted hairpin turns are shown in Figure IIIB.9 and IIIB.10, respectively.<sup>31,32</sup> Enantiopurities of precursors were probed via analytical chiral HPLC assays of their tryptamide derivatives (data not shown).



**Figure IIIB.9.** Synthetic procedures for the generation of fluorine-substituted hairpin turns.

Reagents and conditions: (a) phthalic anhydride, 140 °C, 0.5 h; (b) 1M HCl, MeOH, Et<sub>2</sub>O, r.t., 3 h, 76% over 2 steps; (c) DAST, pyridine, DCM, 0 °C to r.t., 16 h, 44%; (d) 6M HCl, H<sub>2</sub>O, reflux, 16 h, 91%; (e) Boc<sub>2</sub>O, NaHCO<sub>3</sub>, THF, H<sub>2</sub>O, 0 °C to r.t., 16 h, 71%; (f) p-NO<sub>2</sub>-C<sub>6</sub>H<sub>4</sub>CO<sub>2</sub>H, PPh<sub>3</sub>, DIAD, THF, 0 °C to r.t., 4 h, 93%; (g) Et<sub>3</sub>N, MeOH, r.t., 4 h, 95%; (h) DAST, pyridine, DCM 0 °C to r.t., 16 h, 41%; (i) 6M HCl, H<sub>2</sub>O, reflux, 16 h, 99%; (j) Boc<sub>2</sub>O, NaHCO<sub>3</sub>, THF, H<sub>2</sub>O, 0 °C to r.t., 16 h, 70%.



**Figure IIIB.10.** Synthetic procedures for the generation of hydroxyl-substituted hairpin turns.

Reagents and conditions: (a) Boc<sub>2</sub>O, NaHCO<sub>3</sub>, 1,4-dioxane, H<sub>2</sub>O, 0 °C to r.t., 16 h, 82%; (b) MeI, K<sub>2</sub>CO<sub>3</sub>, DMF, r.t., 44 h, 57%; (c) p-NO<sub>2</sub>-C<sub>6</sub>H<sub>4</sub>CO<sub>2</sub>H, PPh<sub>3</sub>, DIAD, THF, 0 °C to r.t., 8 h, 84%; (d) Et<sub>3</sub>N, MeOH, r.t., 1.5 h, 99%; (e) 1M NaOH, MeOH, H<sub>2</sub>O, 0 °C to r.t., 0.5 h, then 1M HCl, H<sub>2</sub>O, 0 °C, 93%.

(*S*)-(-)-4-amino-2-hydroxybutyric acid (**14**), phthalic anhydride, solid sodium hydrogensulfonate ( $\text{NaHSO}_4$ ), 2M HCl in diethyl ether ( $\text{Et}_2\text{O}$ ), pyridine, diethylaminosulfur trifluoride (DAST), di-*tert*-butyldicarbonate ( $\text{Boc}_2\text{O}$ ), triphenyl phosphine ( $\text{PPh}_3$ ), 4-nitrobenzoic acid, diisopropylazodicarboxylate (DIAD), triethylamine, methyl iodide ( $\text{CH}_3\text{I}$ ), 1M aqueous HCl, anhydrous *N,N*-dimethylformamide (DMF) and tryptamine were purchased from Sigma-Aldrich. 1M aqueous NaOH, concentrated HCl, solid sodium sulfate ( $\text{Na}_2\text{SO}_4$ ), and solid sodium bicarbonate ( $\text{NaHCO}_3$ ) were purchased from EMD Chemicals. Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) was purchased from Novabiochem. Solid citric acid hexahydrate was purchased from Mallinckrodt Chemicals. All NMR spectroscopy data was obtained on a Variant instrument operating at 500 MHz. HRMS were acquired using an Agilent 6200 Series TOF with an Agilent G1978A Multimode source in electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI) or mixed (MM) ionization mode. Optical rotation measurements were measured on a Jasco P-1010 polarimeter at 589 nm in spectrophotometric grade solvents.

**(*S*)-methyl 4-(1,3-dioxoisindolin-2-yl)-2-hydroxy-butanoate (**15**)**

(*S*)-(-)-4-amino-2-hydroxybutyric acid (4.00 g, 33.6 mmol) **14** was combined with phthalic anhydride (4.97 g, 33.6 mmol). The solids were ground, mixed, and heated to 140 °C, at which point the phthalic anhydride melted and dissolved **14**. After 30 min, the reaction mixture was cooled to room temperature and re-dissolved in EtOAc (300mL). The solution was extracted with 1M  $\text{NaHSO}_4$  (100 mL, pH 3), washed with water (3 x 100 mL), and brine (100 mL). The organic layer was dried over  $\text{Na}_2\text{SO}_4$ , filtered, and concentrated in vacuo. To the crude mixture (8.3 g) dissolved in anhydrous methanol (250 mL) was slowly added 2M HCl in  $\text{Et}_2\text{O}$  (250 mL). The reaction was stirred at room temperature; TLC indicated completion after 3 h. The mixture was concentrated on rotovap and purified on a short silica gel column (9:1 DCM:EtOAc). The desired fractions were combined,

concentrated and dried to give product **15** as a fine white powder (6.68 g, 76% yield).  $^1\text{H}$  NMR [499.8 MHz, DMSO- $d_6$ ]:  $\delta$  7.83 (m, 4H), 5.57 (d,  $J = 5.5$  Hz, 1H), 4.11 (m, 1H), 3.68 (t,  $J = 7.25$  Hz, 2H), 3.59 (s, 3H), 2.01 (m, 1H), 1.84 (m, 1H);  $^{13}\text{C}$  NMR [125.7 MHz, DMSO- $d_6$ ]:  $\delta$  173.9, 167.8, 134.3, 131.7, 122.9, 68.0, 51.5, 34.3, 32.2; HRMS (MM: ESI–APCI) calcd for  $\text{C}_{13}\text{H}_{13}\text{NO}_5$   $[\text{M}+\text{Na}]^+$  286.0686, found 286.0684;  $[\alpha]_D^{22} +10.40$  ( $c = 0.5$ ,  $\text{CHCl}_3$ )

**(R)-methyl 4-(1,3-dioxoisindolin-2-yl)-2-fluorobutanoate (16)**

To hydroxy ester **15** (5.00 g, 19.0 mmol) in DCM (50 mL) was added pyridine (1.923 mL, 23.7 mmol) at 0 °C. The solution was purged with argon and DAST (3.11 mL, 23.7 mmol) was added drop-wise under argon pressure. The reaction was stirred at room temperature for 18 h and poured slowly into a cold saturated  $\text{NaHCO}_3$  solution. The organic layer was extracted, washed with water (50 mL), 0.1M HCl (50 mL), water (50 mL) and brine (50 mL), and dried ( $\text{Na}_2\text{SO}_4$ ), filtered, and concentrated in vacuo. The product was purified via silica gel column chromatography (100% DCM to 9:1 DCM: EtOAc). The desired fractions were combined, concentrated and dried to yield compound **16** as a pale yellow solid (2.21 g, 44% yield).  $^1\text{H}$  NMR [499.8 MHz, DMSO- $d_6$ ]:  $\delta$  7.85 (m, 4H), 5.24 (ddd,  $J_{\text{H-H}} = 3.5$  Hz,  $J_{\text{H-H}} = 8.5$  Hz,  $J_{\text{H-F}} = 48$  Hz, 2H), 3.72 (m, 2H), 3.65 (s, 3H), 2.09-2.29 (brm, 2H);  $^{13}\text{C}$  NMR [125.7 MHz, DMSO- $d_6$ ]:  $\delta$  169.3 ( $^2J_{\text{C-F}} = 23.5$  Hz), 167.7, 134.3, 131.7, 123.0, 86.7 ( $^1J_{\text{C-F}} = 181.0$  Hz), 52.2, 33.1 ( $^3J_{\text{C-F}} = 3.3$  Hz), 30.3 ( $^2J_{\text{C-F}} = 20.7$  Hz); HRMS (MM: ESI–APCI) calcd for  $\text{C}_{13}\text{H}_{12}\text{FNO}_4$   $[\text{M}+\text{H}]^+$  266.0823, found 266.0830;  $[\alpha]_D^{22} -2.64$  ( $c = 0.5$ ,  $\text{CHCl}_3$ )

**(R)-4-amino-2-fluorobutanoic acid (17)**

Ester **16** (300 mg, 1.13 mmol) was suspended in 6M HCl (3.0 mL) and refluxed for 16 h under argon pressure. The mixture was cooled to room temperature, the phthalic acid precipitate was removed by filtration. The filtrate was concentrated in vacuo, and the

product was purified by recrystallization in iPrOH and Et<sub>2</sub>O to afford **17** as white crystals (124mg, 91% yield). <sup>1</sup>H NMR [499.8 MHz, D<sub>2</sub>O]: δ 5.08 (ddd,  $J_{\text{H-H}} = 4.0$  Hz,  $J_{\text{H-H}} = 8.0$  Hz,  $J_{\text{H-F}} = 49.0$  Hz, 2H), 3.10 (m, 2H), 2.09-2.32 (brm, 2H); <sup>13</sup>C NMR [125.7 MHz, D<sub>2</sub>O]: δ 173.4 ( $^2J_{\text{C-F}} = 23.0$  Hz), 87.6 ( $^1J_{\text{C-F}} = 180.0$  Hz), 36.0 ( $^3J_{\text{C-F}} = 3.8$  Hz), 29.6 ( $^2J_{\text{C-F}} = 20.7$  Hz); HRMS (MM: ESI-APCI) calc'd for C<sub>4</sub>H<sub>8</sub>FNO<sub>4</sub> [M+H]<sup>+</sup> 122.0612, found 122.061; [α]<sub>D</sub><sup>22</sup> +10.64 (c = 0.5, H<sub>2</sub>O)

**(R)-4-(tert-butoxycarbonylamino)-2-fluorobutanoic acid (18)**

To a solution of amino acid **17** (40 mg, 0.33 mmol) in THF (1 mL) and water (1 mL), solid NaHCO<sub>3</sub> (83 mg, 0.99 mmol) and Boc<sub>2</sub>O (87 mg, 0.40 mmol) was added at 0 °C. After 30 min, the solution was slowly warmed to room temperature and stirred for 16 h, followed by extraction with Et<sub>2</sub>O (2 x 1.5mL) to remove generated *tert*-butanol and unreacted Boc<sub>2</sub>O. The aqueous layer was carefully acidified to pH 2.5 with half-saturated citric acid at 0 °C, and extracted with DCM (3 x 1.5 mL). The combined organic phase was dried (Na<sub>2</sub>SO<sub>4</sub>), filtered, concentrated and dried to give product **18** as a colorless oil (52 mg, 71% yield). <sup>1</sup>H NMR [499.8 MHz, DMSO-d<sub>6</sub>]: δ 13.4 (brs, 1H), 6.91 (t,  $J = 5.5$  Hz, 1H), 4.95 (ddd,  $J_{\text{H-H}} = 3.5$  Hz,  $J_{\text{H-H}} = 8.5$  Hz,  $J_{\text{H-F}} = 49.0$  Hz, 2H), 3.04 (m, 2H), 1.75-2.01 (brm, 2H), 1.37 (s, 9H); <sup>13</sup>C NMR [125.7 MHz, DMSO-d<sub>6</sub>]: δ 170.9 ( $^2J_{\text{C-F}} = 23.0$  Hz), 155.6, 86.6 ( $^1J_{\text{C-F}} = 180.5$  Hz), 77.7, 35.7, 32.4 ( $^2J_{\text{C-F}} = 20.2$  Hz), 28.2; HRMS (MM: ESI-APCI) calcd for C<sub>9</sub>H<sub>16</sub>FNO<sub>4</sub> [M-H]<sup>-</sup> 220.0991, found 220.1018; [α]<sub>D</sub><sup>22</sup> +5.32 (c = 0.5, MeOH)

**(R)-methyl 4-(1,3-dioxoisindolin-2-yl)-2-hydroxy-butanoate (19)**

To a solution of **15** (5.00g, 19.0 mmol) in THF (125 mL) was added PPh<sub>3</sub> (5.485 g, 20.9 mmol) followed by 4-nitrobenzoic acid (9.53 g, 57.0 mmol). DIAD (4.06 mL, 20.9 mmol) was added dropwise to the solution at 0 °C, and the reaction was stirred at room temperature. TLC showed reaction completion after 4 h, and the solution was concentrated in vacuo. The product was isolated via silica gel column chromatography (100% DCM),

and the combined fractions were concentrated and dried in vacuo to afford the Mitsunobu intermediate (*R*)-4-(1,3-dioxoisindolin-2-yl)-1-methoxy-1-oxobutan-2-yl 4-nitrobenzoate as a pale yellow powder (7.3 g, 93% yield).  $^1\text{H}$  NMR [499.8 MHz, DMSO- $d_6$ ]:  $\delta$  8.30 (dd,  $J = 2.5$  Hz,  $J = 9$  Hz, 2H), 8.17 (dd,  $J = 2.5$  Hz,  $J = 9$  Hz, 2H), 7.78 (m, 4H), 5.33 (dd,  $J = 5$  Hz,  $J = 7.5$  Hz, 1H), 3.84 (m, 2H), 3.65 (s, 3H), 2.33 (m, 2H);  $^{13}\text{C}$  NMR [125.7 MHz, DMSO- $d_6$ ]:  $\delta$  169.4, 167.9, 163.6, 150.4, 134.3, 134.0, 131.6, 130.7, 123.9, 123.0, 71.2, 52.5, 33.9, 30.0; HRMS (MM: ESI-APCI) calcd for  $\text{C}_{20}\text{H}_{16}\text{N}_2\text{O}_8$   $[\text{M}+\text{H}]^+$  413.0979, found 413.0975;  $[\alpha]^{22}_{\text{D}} +28.6$  ( $c = 0.5$ ,  $\text{CHCl}_3$ ).

To obtain the desired product **19**, triethylamine (9.93 mL, 77.5 mmol) was added to the intermediate (6.4 g, 15.5 mmol) in anhydrous methanol (480 mL). The reaction was monitored by TLC and was complete after 4 h. The solution was then concentrated and purified via column chromatography (100% DCM to 9:1 DCM:EtOAc). The desired fractions were combined, concentrated and dried to afford compound **19** as a fine white powder (3.86 g, 95% yield).  $^1\text{H}$  NMR [499.8 MHz, DMSO- $d_6$ ]:  $\delta$  7.84 (m, 4H), 5.57 (d,  $J = 5.5$  Hz, 1H), 4.12 (m, 1H), 3.68 (t,  $J = 7.5$  Hz, 2H), 3.59 (s, 3H), 2.02 (m, 1H), 1.84 (m, 1H);  $^{13}\text{C}$  NMR [125.7 MHz, DMSO- $d_6$ ]:  $\delta$  173.9, 167.8, 134.3, 131.7, 122.9, 67.9, 51.5, 34.3, 32.1; HRMS (MM: ESI-APCI) calcd for  $\text{C}_{13}\text{H}_{13}\text{NO}_5$   $[\text{M}+\text{Na}]^+$  286.0686, found 286.0688  $[\alpha]^{22}_{\text{D}} -10.28$  ( $c = 0.5$ ,  $\text{CHCl}_3$ ).

### **(*S*)-methyl 4-(1,3-dioxoisindolin-2-yl)-2-fluorobutanoate (20)**

Synthesis was performed using *R*-enantiomer **19** (3.85 g, 14.6 mmol) in conjunction with the method for the synthesis of fluoroester **16**. The product **20** was synthesized and isolated as a pale yellow solid (1.60 g, 41% yield).  $^1\text{H}$  NMR [499.8 MHz, DMSO- $d_6$ ]:  $\delta$  7.85 (m, 4H), 5.24 (ddd,  $J_{\text{H-H}} = 3.5$  Hz,  $J_{\text{H-H}} = 8.5$  Hz,  $J_{\text{H-F}} = 48.0$  Hz, 2H), 3.72 (m, 2H), 3.65 (s, 3H), 2.09-2.29 (brm, 2H);  $^{13}\text{C}$  NMR [125.7 MHz, DMSO- $d_6$ ]:  $\delta$  169.3 ( $^2J_{\text{C-F}} = 23.5$  Hz), 167.7, 134.3, 131.7, 123.0, 86.7 ( $^1J_{\text{C-F}} = 181.3$  Hz), 52.2, 33.1 ( $^3J_{\text{C-F}} = 3.8$  Hz), 30.3 ( $^2J_{\text{C-F}} = 20.2$  Hz); HRMS (MM: ESI-APCI) calcd for  $\text{C}_{13}\text{H}_{12}\text{FNO}_4$   $[\text{M}+\text{H}]^+$  266.0823, found



266.0828;  $[\alpha]_D^{22} +2.4$  ( $c = 0.5$ ,  $\text{CHCl}_3$ ).

**(S)-4-amino-2-fluorobutanoic acid (21)**

This compound was synthesized by following the method used for the synthesis of amino acid **17**, but using the *S*-enantiomer **20** (1.60 g, 6.03 mmol) as starting material. The product was synthesized and isolated as fluffy white crystals (725 mg, 99% yield).  $^1\text{H}$  NMR [499.8 MHz,  $\text{D}_2\text{O}$ ]:  $\delta$  5.08 (ddd,  $J_{\text{H-H}} = 4.0$  Hz,  $J_{\text{H-H}} = 8.0$  Hz,  $J_{\text{H-F}} = 49.0$  Hz, 2H), 3.09 (m, 2H), 2.08-2.32 (brm, 2H);  $^{13}\text{C}$  NMR [125.7 MHz,  $\text{D}_2\text{O}$ ]:  $\delta$  173.4 ( $^2J_{\text{C-F}} = 23.0$  Hz), 87.7 ( $^1J_{\text{C-F}} = 180.0$  Hz), 36.0 ( $^3J_{\text{C-F}} = 3.3$  Hz), 29.6 ( $^2J_{\text{C-F}} = 20.7$  Hz); HRMS (MM: ESI-APCI) calcd for  $\text{C}_4\text{H}_8\text{FNO}_4$   $[\text{M}+\text{H}]^+$  122.0612, found 122.0613;  $[\alpha]_D^{22} -11.2$  ( $c = 0.5$ ,  $\text{H}_2\text{O}$ ).

**(S)-4-(tert-butoxycarbonylamino)-2-fluorobutanoic acid (22)**

This compound was synthesized by following the method used for the synthesis of amino acid **18**, but using the *S*-enantiomer **21** (40 mg, 0.33 mmol) as starting material. The product **22** was synthesized and isolated as a colorless oil (51 mg, 70% yield).  $^1\text{H}$  NMR [499.8 MHz,  $\text{DMSO-d}_6$ ]:  $\delta$  13.3 (brs, 1H), 6.91 (t,  $J = 5.5$  Hz, 1H), 4.95 (ddd,  $J_{\text{H-H}} = 3.5$  Hz,  $J_{\text{H-H}} = 8.5$  Hz,  $J_{\text{H-F}} = 49.0$  Hz, 2H), 3.04 (m, 2H), 1.75-2.01 (brm, 2H), 1.37 (s, 9H);  $^{13}\text{C}$  NMR [125.7 MHz,  $\text{DMSO-d}_6$ ]:  $\delta$  170.8 ( $^2J_{\text{C-F}} = 23.5$  Hz), 155.4, 86.4 ( $^1J_{\text{C-F}} = 180.5$  Hz), 77.6, 35.6, 32.3 ( $^2J_{\text{C-F}} = 19.9$  Hz), 28.1; HRMS (MM: ESI-APCI) calcd for  $\text{C}_9\text{H}_{16}\text{FNO}_4$   $[\text{M}-\text{H}]^-$  220.0991, found 220.1017;  $[\alpha]_D^{22} -5.04$  ( $c = 0.5$ ,  $\text{MeOH}$ ).

**(S)-4-(tert-butoxycarbonylamino)-2-hydroxybutanoic acid (23)**

To a solution of (*S*)-(-)-4-amino-2-hydroxybutyric acid **14** (2.00 g, 16.8 mmol) dissolved in water (50 mL) and 1,4-dioxane (50 mL),  $\text{NaHCO}_3$  (1.55 g, 18.5 mmol) and added  $\text{Boc}_2\text{O}$  (4.04 g, 18.5 mmol) was added at 0 °C. After 15 min, the turbid solution was slowly warmed to room temperature and stirred for 16 h. The solution was extracted with  $\text{Et}_2\text{O}$  (2 x 50 mL) to rid of the generated *tert*-butanol and unreacted di-*tert*-butyldicarbonate.

The aqueous layer was carefully acidified to pH 2.5 with 1M HCl (~20 mL) at 0 °C, and extracted with DCM (3 x 50 mL). The combined organic phase was dried (Na<sub>2</sub>SO<sub>4</sub>), filtered, concentrated and dried to give product **23** as a colorless oil (3.03 g, 82% yield). <sup>1</sup>H NMR [499.8 MHz, DMSO-d<sub>6</sub>]: δ 6.75 (t, *J* = 5.0 Hz, 1H), 3.93 (dd, *J* = 4.0 Hz, *J* = 8.5 Hz, 1H), 3.00 (m, 2H), 1.77 (m, 1H), 1.57 (m, 1H), 1.36 (s, 9H); <sup>13</sup>C NMR [125.7 MHz, DMSO-d<sub>6</sub>]: δ 175.6, 155.5, 77.5, 67.7, 36.7, 34.1, 28.2; HRMS (MM: ESI-APCI) calcd for C<sub>9</sub>H<sub>17</sub>NO<sub>5</sub> [M+Na]<sup>+</sup> 242.0999, found 242.0997; [α]<sub>D</sub><sup>22</sup> -3.96 (c = 0.5, MeOH).

**(S)-methyl 4-(tert-butoxycarbonylamino)-2-hydroxybutanoate (24)**

To a solution of Boc-protected amino acid **23** (2.00 g, 9.1 mmol) dissolved in DMF (20 mL) was added K<sub>2</sub>CO<sub>3</sub> (1.32 g, 9.54 mmol) at 0 °C. After 15 min, CH<sub>3</sub>I (4.13 g, 29.12 mmol) was added slowly and the solution was stirred for 44 h at room temperature. The solution was then partitioned between H<sub>2</sub>O (100mL) and ethyl acetate (4 x 50 mL). The organic fractions were combined, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered, concentrated and dried to give product **24** as a pale yellow oil (1.20 g, 57% yield). <sup>1</sup>H NMR [499.8 MHz, DMSO-d<sub>6</sub>]: δ 6.76 (t, *J* = 5.0 Hz, 1H), 5.39 (d, *J* = 6.0 Hz, 1H), 4.04 (m, 1H), 3.62 (s, 3H), 2.99 (m, 2H), 1.75 (m, 1H), 1.59 (m, 1H), 1.36 (s, 9H); <sup>13</sup>C NMR [125.7 MHz, DMSO-d<sub>6</sub>]: δ 174.4, 155.5, 77.5, 67.8, 51.4, 36.4, 34.1, 28.2; HRMS (MM: ESI-APCI) calcd for C<sub>10</sub>H<sub>19</sub>NO<sub>5</sub> [M+Na]<sup>+</sup> 256.1155, found 256.1159; [α]<sub>D</sub><sup>22</sup> +4.56 (c = 0.5, MeOH).

**(R)-methyl 4-(tert-butoxycarbonylamino)-2-hydroxybutanoate (25)**

To a solution of ester **24** (650 mg, 2.75 mmol) in THF (13 mL) was added PPh<sub>3</sub> (795 g, 3.02 mmol) followed by 4-nitrobenzoic acid (1.38 g, 8.26 mmol). DIAD (589 mL, 3.03 mmol) was added dropwise to the solution at 0 °C, and the reaction was stirred at room temperature. TLC showed a completed reaction after 8 h, and the solution was concentrated in vacuo. The product was isolated via silica gel column chromatography (95:5 DCM: EtOAc to 9:1 DCM:EtOAc), and the combined fractions were concentrated and dried in

vacuo to afford the Mitsunobu intermediate (*R*)-4-(*tert*-butoxycarbonylamino)-1-methoxy-1-oxobutan-2-yl 4-nitrobenzoate as a colorless oil (890 mg, 84% yield). <sup>1</sup>H NMR [499.8 MHz, DMSO-d<sub>6</sub>]: δ 8.38 (dd, *J* = 2.0 Hz, *J* = 9.0 Hz, 2H), 8.24 (dd, *J* = 2.0 Hz, *J* = 9.0 Hz, 2H), 6.99 (t, 5.5 Hz, 1H), 5.20 (m, 1H), 3.70 (s, 3H), 3.16 (m, 2H), 2.05 (m, 2H), 1.34 (s, 9H); <sup>13</sup>C NMR [125.7 MHz, DMSO-d<sub>6</sub>]: δ 169.9, 163.8, 155.6, 150.5, 134.3, 130.8, 124.0, 77.7, 70.9, 52.4, 35.8, 30.8, 28.1; HRMS (MM: ESI-APCI) calcd for C<sub>17</sub>H<sub>2</sub>N<sub>2</sub>O<sub>8</sub> [M-Boc+2H]<sup>+</sup> 283.0925, found 283.0934; [α]<sub>D</sub><sup>22</sup> +16.64 (c = 0.5, MeOH).

To obtain the desired product **25**, triethylamine (750 mL, 5.38 mmol) was added to the intermediate (410 mg, 1.07 mmol) in anhydrous methanol (30 mL). The reaction was monitored by TLC and was complete after 1.5 h. The solution was then concentrated and run through a silica plug. DCM was run till the eluent was no longer UV active, at which point EtOAc was used to flush out the product. The EtOAc fractions were combined, concentrated and dried to afford compound **25** as a pale yellow oil (250 mg, 99% yield). <sup>1</sup>H NMR [499.8 MHz, DMSO-d<sub>6</sub>]: δ 6.76 (t, *J* = 5.0 Hz, 1H), 5.39 (d, *J* = 6.0 Hz, 1H), 4.04 (m, 1H), 3.62 (s, 3H), 2.99 (m, 2H), 1.76 (m, 1H), 1.60 (m, 1H), 1.36 (s, 9H); <sup>13</sup>C NMR [125.7 MHz, DMSO-d<sub>6</sub>]: δ 174.4, 155.5, 77.5, 67.8, 51.4, 36.5, 34.1, 28.2; HRMS (MM: ESI-APCI) calcd for C<sub>10</sub>H<sub>19</sub>NO<sub>5</sub> [M+Na]<sup>+</sup> 256.1155, found 256.1158; [α]<sub>D</sub><sup>22</sup> -4.28 (c = 0.5, MeOH).

#### (*R*)-4-(*tert*-butoxycarbonylamino)-2-hydroxybutanoic acid (**26**)

To a solution of Boc-protected amino acid **25** in MeOH (2.5 mL) was added dropwise 1M NaOH (2.5 mL) at 0 °C. The reaction mixture was warmed to room temperature and TLC showed a completed reaction after 30 min. The solution was cooled to 0 °C and re-acidified to pH2 by dropwise addition of 1M HCl (2.5 mL), and extracted with EtOAc (3 x 2.5 mL). The organic fractions were then combined, dried (Na<sub>2</sub>SO<sub>4</sub>), filtered, concentrated and dried to give product **26** as a colorless oil (108 mg, 93% yield). <sup>1</sup>H NMR [499.8 MHz, DMSO-d<sub>6</sub>]: δ 6.75 (t, *J* = 5.0 Hz, 1H), 3.92 (dd, *J* = 4.0 Hz, *J* = 8.5 Hz, 1H), 3.00 (m, 2H),

169  
1.77 (m, 1H), 1.56 (m, 1H), 1.36 (s, 9H);  $^{13}\text{C}$  NMR [125.7 MHz, DMSO- $d_6$ ]:  $\delta$  175.6, 155.5, 77.4, 67.8, 36.7, 34.1, 28.2; HRMS (MM: ESI–APCI) calcd for  $\text{C}_9\text{H}_{17}\text{NO}_5$   $[\text{M}+\text{Na}]^+$  242.0999, found 242.0998;  $[\alpha]_D^{22} +4.12$  (c = 0.5, MeOH).

## Acknowledgements

We are grateful to the National Institutes of Health for research support. We thank David M. Chenoweth for helpful discussions. Mass spectrometry analyses were performed in the Mass Spectrometry Facility of the Division of Chemistry and Chemical Engineering at the California Institute of Technology.

## References

1. Dervan, P. B. *Bioorg. Med. Chem.* **2001**. *9*, 2215-2235.
2. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**. *13*, 284-299.
3. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003**. *100*, 12063-12068.
4. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004**. *32*, 2802-2818.
5. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2004**. *101*, 16768-16773.
6. Kageyama, Y.; Sugiyama, H.; Ayame, H.; Iwai, A.; Fujii, Y.; Huang, L. E.; Kizaka-Kondoh, S.; Hiraoka, M.; Kihara, K. *Acta Oncol.* **2006**. *45*, 317-324.
7. Fukuda, N.; Ueno, T.; Tahira, Y.; Ayame, H.; Zhang, W.; Bando, T.; Sugiyama, H.; Saito, S.; Matsumoto, K.; Mugishima, H.; Serie, K. *J. Am. Soc. Nephrol.* **2006**. *17*, 422-432.
8. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**. *35*, 363-370.
9. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *ACS Chem. Biol.* **2007**. *2*, 561-571.
10. Matsuda, H.; Fukuda, N.; Yao, E. H.; Ueno, T.; Sugiyama, H.; Matsumoto, K. *J. Hypertens.* **2008**. *26*, S197-S197.
11. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C. *Nat. Struct. Biol.* **1998**. *5*, 104-109.
12. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature*. **1998**. *391*, 468-471.
13. Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U. S. A.* **1989**. *86*, 5723-5727.
14. Swalley, S. E.; Baird, E. E.; Dervan, P. B.; *J. Am. Chem. Soc.* **1999**. *121*, 1113-1120.
15. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**. *116*, 7983-7988.
16. Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**. *120*, 1382-1391.

17. Zhang, W.; Minoshima, M.; Sugiyama, H. *J. Am. Chem. Soc.* **2006.** *128*, 14905-14912.
18. Farkas, M. E.; Tsai, S. M.; Dervan, P. B. *Bioorg. Med. Chem.* **2007.** *15*, 6927-6936.
19. Gellman, S. H.; Adams, B. R.; Dado, G. P. *J. Am. Chem. Soc.* **1990.** *112*, 460-461.
20. Appella, D. H.; Christianson, L. A.; Klein, D. A.; Powell, D. R.; Huang, X. L.; Barchi, J. J.; Gellman, S. H. *Nature.* **1997.** *387*, 381-384.
21. Gellman, S. H. *Accounts Chem. Res.* **1998.** *31*, 173-180.
22. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B. Z.; Puckett, J. W.; Wang, C. C. C.; Dervan, P. B. *Tetrahedron* **2007.** *63*, 6146-6151.
23. Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B. *J. Am. Chem. Soc.* **2008.** *130*, 6859-6866.
24. Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Method Enzymol.* **1986.** *130*, 132-181.
25. Senear, D. F.; Dalmawiszhausz, D. D.; Brenowitz, M. *Electrophoresis.* **1993.** *14*, 704-712.
26. Trauger, J. W.; Dervan, P. B. *Meth. Enzymol.* **2001.** *340*, 450-466.
27. Pilch, D. S.; Poklar, N.; Gelfand, C. A.; Law, S. M.; Breslauer, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **1996.** *93*, 8306-8311.
28. Pilch, D. S.; Poklar, N.; Baird, E. E.; Dervan, P. B.; Breslauer, K. J. *Biochemistry.* **1999.** *38*, 2143-2151.
29. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature.* **1996.** *382*, 559-561.
30. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996.** *118*, 6141-6146.
31. Busby, G. W. *Ph.D. Thesis, Harvard University.* **1974.** 48-54.
32. Hoshi, H.; Aburaki, S.; Iimura, S.; Yamasaki, T.; Naito, T.; Kawaguchi, H. *J. Antibiot.* **1990.** *43*, 858-872.

## Appendix

### **Exploration of Methods for Genome-wide Analysis of Polyamide Binding**

*Several experiments in this chapter used polyamides synthesized by Vadim Baidin and Daniel A. Harki (California Institute of Technology).*

## Abstract

Pyrrole-imidazole polyamides targeted to bind promoter sequences of DNA have been used to modulate gene expression by preventing the binding of transcriptional activators. We would like to be able to determine the genome-wide binding of polyamides in the context of chromatin *in vivo*. To this end, we have synthesized polyamide-biotin conjugates to pull down DNA-polyamide complexes using streptavidin beads. Polyamides targeting the sequence 5'-WGWWCW-3' located within the androgen response element (ARE) have been shown to down-regulate androgen responsive genes. We have used RT-PCR as a surrogate assay for cellular uptake of biotin-polyamides targeted to the ARE, and shown that our compounds are able to bind DNA and recruit streptavidin *in vitro*. We have also conducted experiments where chromatin from polyamide-treated cells has been harvested, sheared, and pulled-down with streptavidin beads. The amplification of androgen responsive genes and loci containing multiple polyamide match sites has been measured using RT-PCR. It is likely that the currently used method of shearing DNA results in polyamide displacement and subsequent re-equilibration. Future efforts may involve redesign of the polyamide to include a moiety capable of cross-linking to DNA in order to prevent this. The methods described herein should be applicable to these endeavors.

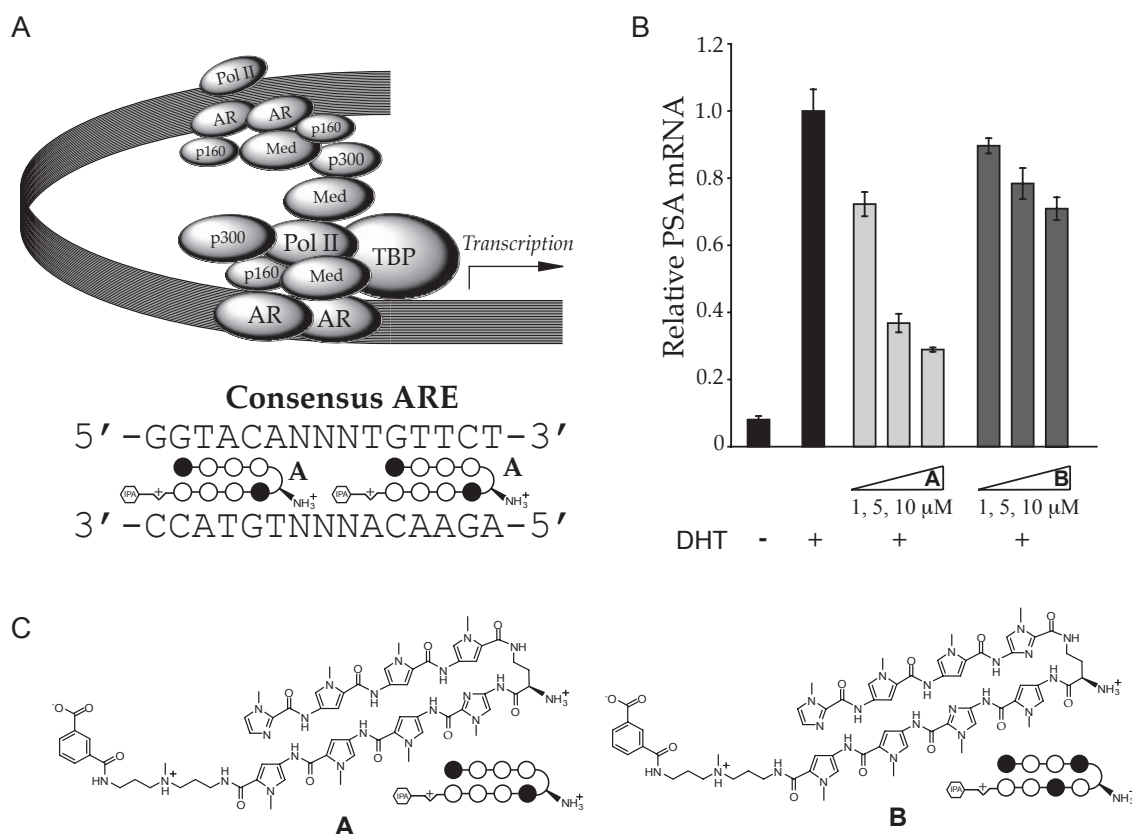


## Introduction

The androgen receptor (AR) is a member of the ligand-activated nuclear receptor family of transcription factors.<sup>1</sup> Ligand binding to AR initiates its release from the cytoplasm followed by dimerization and binding to androgen response elements (ARE) of target genes, which results in gene activation through interaction with co-activators and the general transcription machinery.<sup>2</sup> Functional AREs have the consensus sequence 5'-GGTACAnnnTGTTCT-3'.<sup>3</sup> Work in the Dervan laboratory has previously shown that a polyamide designed to target the sequence 5'-WGWWCW-3', found in the consensus ARE, is cell-permeable, binds the ARE in the prostate specific antigen (PSA) promoter, and inhibits the expression of PSA and other androgen responsive genes in cultured prostate cancer cells (Figure A.1). A control polyamide targeted to the sequence 5'-WGWCGW-3' had little effect. Chromatin immunoprecipitation (ChIP) has shown that match polyamide treatment results in reduced AR occupancy at the PSA promoter and enhancer.<sup>4</sup>

Polyamide binding has previously been studied in several *in vitro* contexts. DNase I footprinting titrations on designed DNA fragments roughly 100 base pairs in size enable the determination of equilibrium association constants and allow for limited exploration of specificities.<sup>5,6</sup> The more recently developed cognate site identifier (CSI) microarray platform, which presents 32,896 unique eight-mers, allows for an unbiased interrogation of polyamide binding preferences.<sup>7,8</sup> We now wish to determine genome-wide profiles for polyamide binding *in vivo* by precipitation of complexes followed by DNA sequencing.

ChIP is a standard method by which the associations of proteins with DNA are determined.<sup>9</sup> Typically, living cells are treated with formaldehyde in order to crosslink proteins to DNA, and then the chromatin is fragmented by sonication. The resulting input sample is immunoprecipitated with an antibody that specifically recognizes a protein of interest; this enriches DNA sequences that associate with the protein *in vivo*. The immunoprecipitation efficiency (IE) for an experiment is the amount of DNA in the IP sample divided by the amount of DNA in the input sample. IP efficiencies for protein-



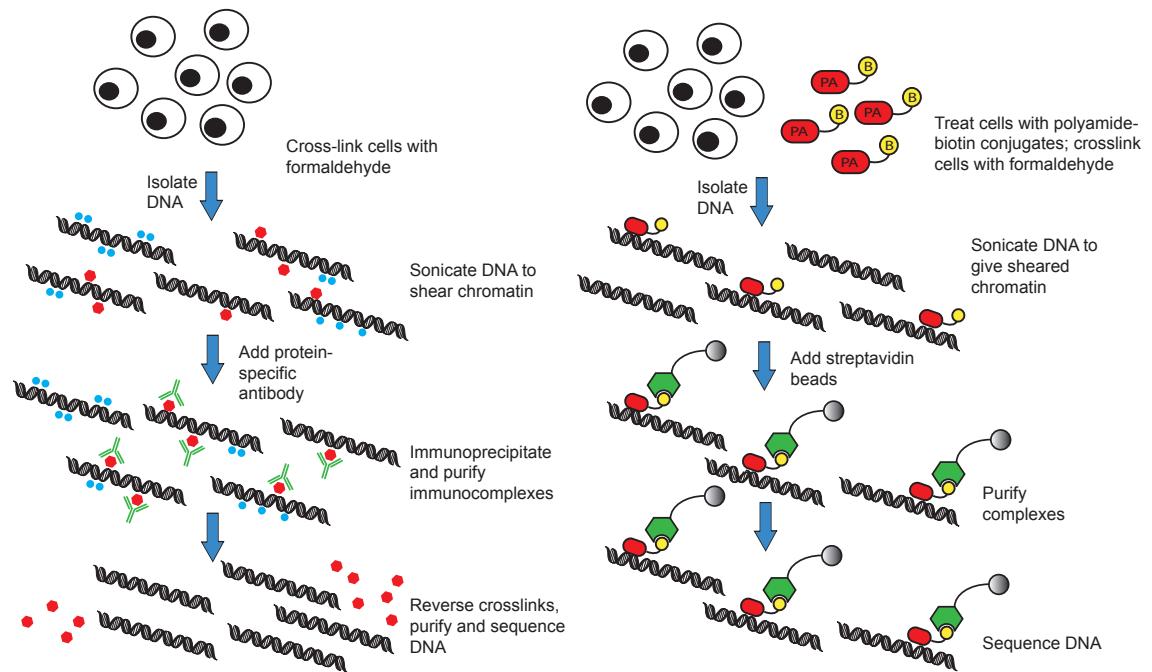
**Figure A.1.** Polyamides targeting the androgen response element. (A) Schematic diagram of the androgen receptor complex (top), and ball-and-stick models of polyamide **A** shown targeting the consensus ARE sequence (bottom). (B) Match polyamide **A** results in a decrease in PSA mRNA, while control polyamide **B** shows little effect following treatment in cell culture. No polyamide control samples are shown in black. (C) Chemical structures and ball-and-stick models of polyamides **A** (left) and **B** (right).<sup>4</sup>

bound regions are higher than for controls, and the relative level of protein binding is described as the fold enrichment over control.

The sensitivity of a ChIP experiment depends on the ability to separate protein-bound fragments of DNA from the unbound fragment background. For this reason, the quality of the antibody is a critical factor. Additionally, a significant concern is the non-specific association of the protein with DNA captured by crosslinking. It has been shown that extensive fragmentation of chromatin significantly increases fold enrichments of protein binding sites in ChIP experiments.<sup>10</sup>

When individual regions are analyzed, amounts of DNA present are typically determined by quantitative PCR in real time (RT-PCR). The resulting material can be analyzed on microarrays,<sup>11</sup> or via high throughput DNA sequencing.<sup>12</sup> ChIP followed by direct ultra high-throughput DNA sequencing of the immunoprecipitated sample (ChIP-Seq) enables binding events to be resolved at 25 base pair resolution and requires no prior knowledge of sequence content.

The affinity of biotin for streptavidin is one of the most stable interactions in biology, with  $K_d = 10^{-15}$ .<sup>13</sup> Streptavidin-coated magnetic particles provide for efficient binding and separation of biotinylated compounds, and have been widely used for the purification of proteins and nucleic acids.<sup>14,15</sup> By utilizing polyamide-biotin conjugates, we plan to pull down polyamide-bound chromatin using streptavidin beads (Figure A.2). We use the polyamide scaffold shown to affect androgen responsive genes in prostate cancer cells as a platform upon which to base these experiments. Procedures common to ChIP serve as a starting-point for the pull-down of polyamides. The work described here includes the synthesis and screening of polyamides by RT-PCR, electrophoretic mobility shift assays (EMSA) to determine polyamides' abilities to recruit streptavidin, and experiments where chromatin from polyamide-dosed cells was harvested, sheared, pulled down, and analyzed for gene amplification by RT-PCR.

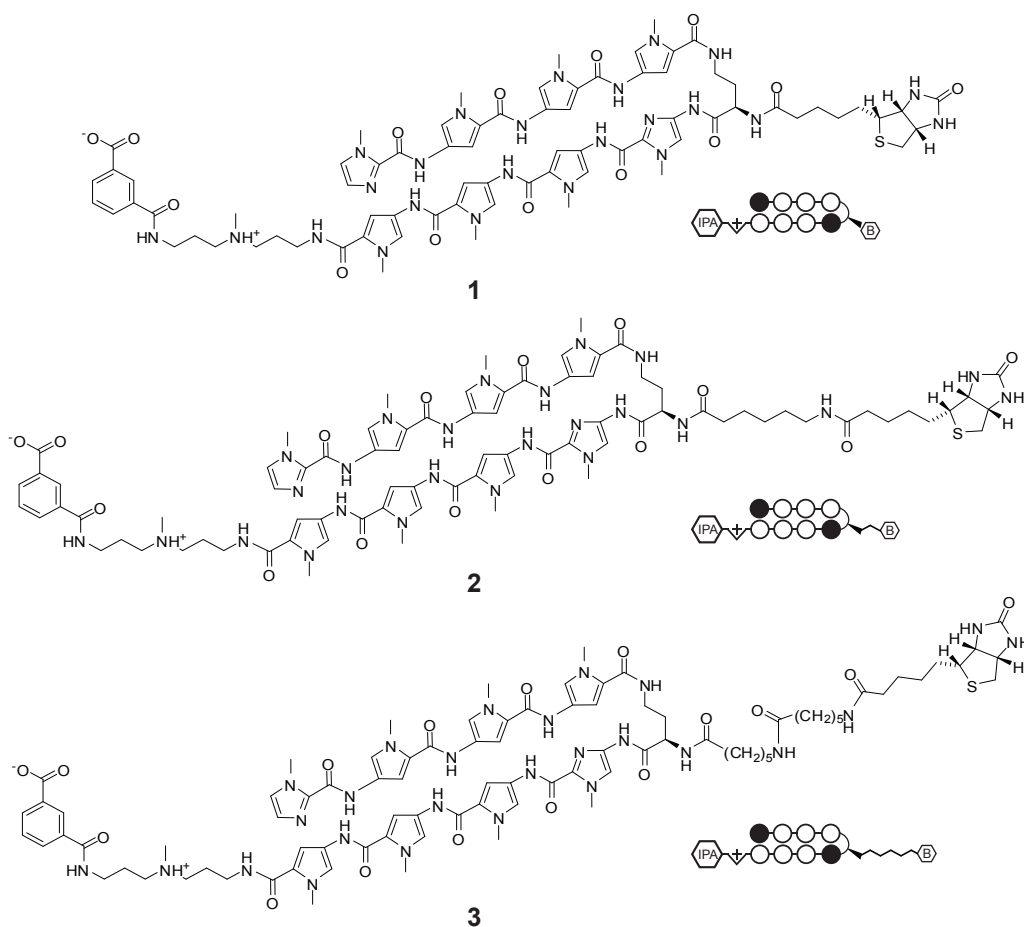


**Figure A.2.** Illustration of chromatin immunoprecipitation versus pull-down with polyamide-biotin conjugates. Major steps for ChIP are shown with double blue circles representing one protein type, and red hexagons indicating another, green 'Y' shapes indicate red-hexagon-specific antibodies (left). Major steps for polyamide pull-down are shown with red hexagons indicating polyamide, yellow circles indicating biotin, green pentagons represent streptavidin, and grey circles represent magnetic beads (right).

## Results and Discussion

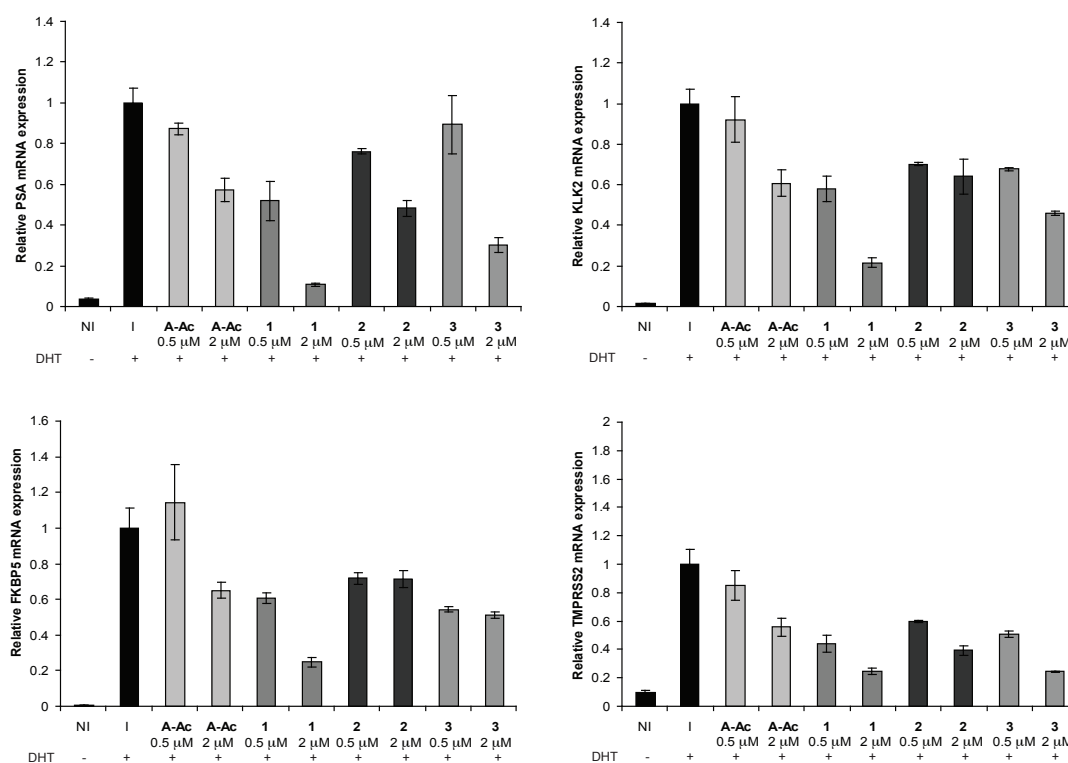
### Polyamides conjugated to biotin at the turn unit

Early efforts to establish biotin-polyamide conjugates as cell-permeable molecules resulted in compounds with minimal effects on gene expression.<sup>16</sup> These compounds contained a biotin moiety at the tail position and targeted the hypoxia response element (HRE, 5'-TACGTG-3'). The presence of an isophthalic acid (IPA) moiety in the tail region has been shown to yield high affinity conjugates with improved nuclear permeability.<sup>16</sup> Subsequently, a series of polyamides targeting the androgen response element (ARE, 5'-WGWWCW-3') were synthesized (Figure A.3). Polyamides **1-3** contain IPA at the tail position, and biotin appended to the turn unit via various linkers.







**Figure A.3.** Structures of polyamides linked to biotin at the turn subunit. Chemical and ball-and-stick structures of polyamides **1-3** are shown. Ball-and-stick representations are as follows: an open circle is pyrrole, a closed circle is imidazole, 'IPA' within a hexagon is isophthalic acid, 'B' within a hexagon is biotin, a semi-circle represents the turn unit.

All three molecules were observed to down-regulate androgen responsive genes in LNCaP cells with **1** and **3** resulting in the greatest effects (Figure A.4). However, both showed decreased binding to the PSA ARE sequence in comparison with polyamides containing acetylated amines at the turn unit (Table A.1). Because hairpin polyamides linked by 3,4-diaminobutyric acid have been shown to possess improved binding affinities and tolerances for modification at the turn unit,<sup>17</sup> analogues of **1** and **3** containing (*R*)-3,4-diaminobutyric acid were prepared. These molecules demonstrated no effect on gene expression (data not shown). Similarly, an analogue of **1** was prepared using PAM resin, resulting in a  $\beta$ -alanine moiety in the tail portion of the polyamide. This molecule showed



**Figure A.4.** Effects of polyamides **1-3** on the expression of androgen responsive genes. LNCaP cells were treated with polyamides **1-3** at 0.5  $\mu$ M and 2  $\mu$ M concentrations. The effects of these polyamides were compared to not induced (NI), induced (I), and A-Ac (the acetylated analogue of polyamide A) treatment conditions. Expression levels for PSA (top, left), KLK2 (top, right), FKBP5 (bottom, left), and TMPRSS2 (bottom, right) are shown. DHT = dihydrotestosterone.

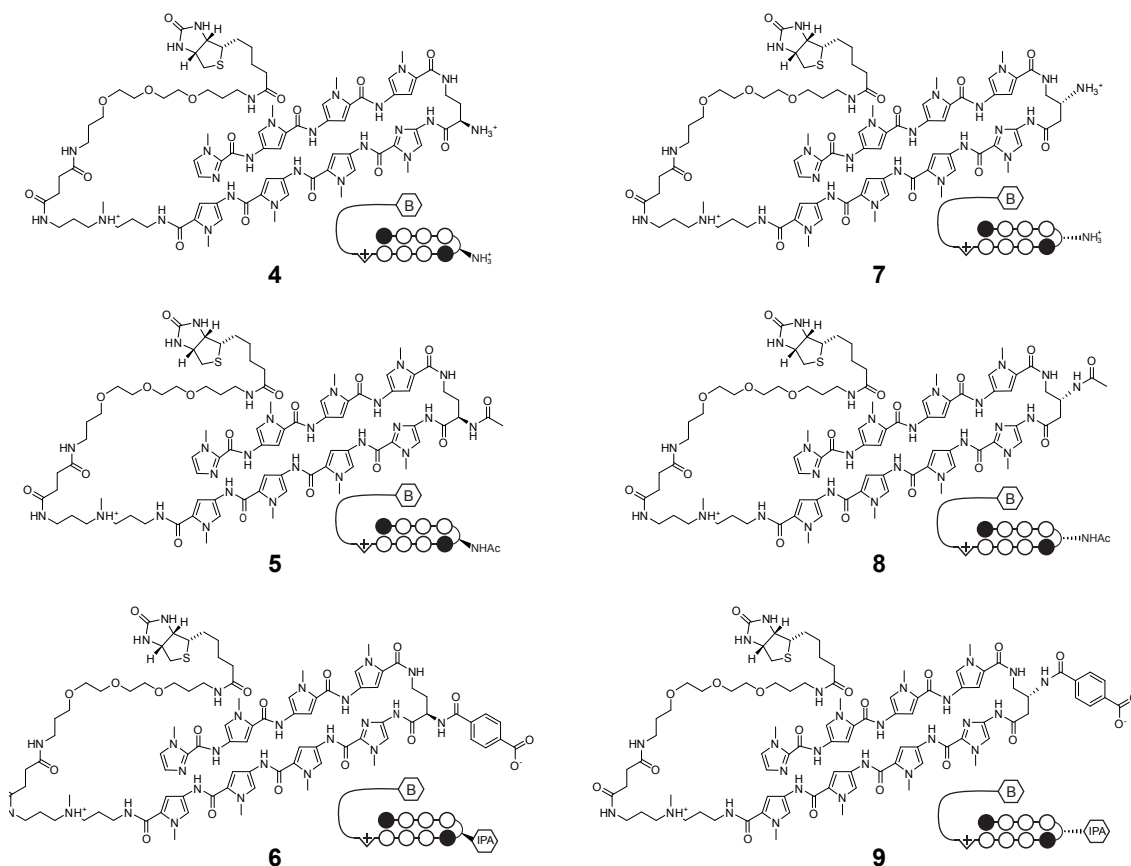
DNA = 5'-TTGC <b>AGAACA</b> GCAA-3'			
Polyamides		$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$
—		60.1 ( $\pm 0.2$ )	—
<b>A-Ac</b>		70.1 ( $\pm 0.2$ )	10.0
<b>C-Ac</b>		74.9 ( $\pm 0.1$ )	14.8
<b>1</b>		69.3 ( $\pm 0.2$ )	9.2
<b>3</b>		59.8 ( $\pm 0.2$ )	5.6

**Table A.1.** Melting temperatures of DNA/polyamide complexes. Polyamides **1** and **3** are compared against **A-Ac** and **C-Ac** (containing (*R*)-3,4-diaminobutyric acid in the turn subunit) in binding to a 14mer containing the PSA ARE sequence. All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses.  $\Delta T_m$  values are given as  $T_m^{(\text{DNA/polyamide})} - T_m^{(\text{DNA})}$ .

moderate effects on gene expression (data not shown). The abilities of polyamides **1** and **3** to access and bind streptavidin while bound to DNA were evaluated by using EMSA. A shift was not observed with either conjugate, leading us to believe that streptavidin was inaccessible for binding to the biotin moieties (data not shown).

### Polyamides conjugated to biotin at the tail unit

Work in the Dervan lab showed that a polyamide-biotin conjugate was capable of binding to specific sequences in DNA nanostructures and recruiting streptavidin. The molecule utilized possessed a polyethylene glycol (PEG) linked biotin moiety in the tail subunit of the polyamide, and was also shown to bind streptavidin via EMSA.<sup>18</sup> This molecule served as inspiration for polyamides **4-9** (Figure A.5). Polyamides **4-6** contain (*R*)-2,4-diaminobutyric acid in the turn unit, whereas **7-9** contain (*R*)-3,4-diaminobutyric acid. In both series, molecules containing a free amine at the turn (**4** and **7**), acetylated amines at the turn (**5** and **8**), and IPA conjugated to the amine (**6** and **9**) were made. Acetylation of the




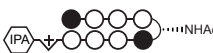



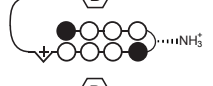
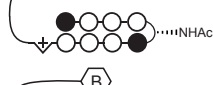

**Figure A.5.** Structures of polyamides linked to biotin at the tail subunit. Chemical and ball-and-stick structures of polyamides **4-6** (containing (*R*)-2,4-diaminobutyric acid in the turn unit (left), and **7-9** (containing (*R*)-3,4-diaminobutyric acid in the turn unit (right)) are shown. Ball-and-stick representations are as follows: an open circle is pyrrole, a closed circle is imidazole, 'IPA' within a hexagon is isophthalic acid, 'B' within a hexagon is biotin, a semi-circle represents the turn unit.

amine has previously been shown to improve cellular uptake of polyamides, and typically results in diminished binding affinities. The presence of IPA in the tail unit also results in improved uptake, however, its conjugation to the turn unit and resulting effects on uptake and affinity had not been previously explored.

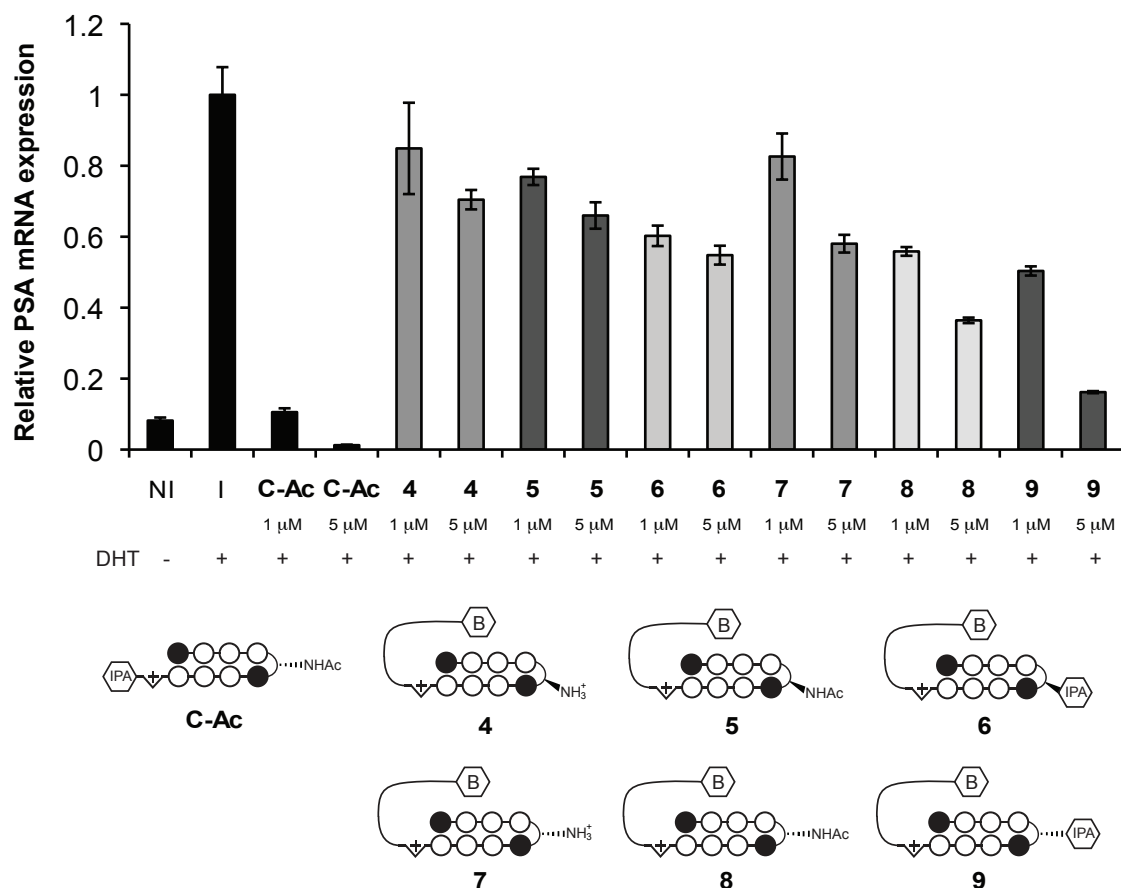
Melting temperature analysis on the PSA ARE revealed that the polyamides containing (*R*)-3,4-diaminobutyric acid had a greater ability to bind DNA than the analogous molecules containing (*R*)-2,4-diaminobutyric acid (Table A.2). Strikingly, the presence of the IPA moieties resulted in only a slight reduction in  $\Delta T_m$ 's compared to the acetylated



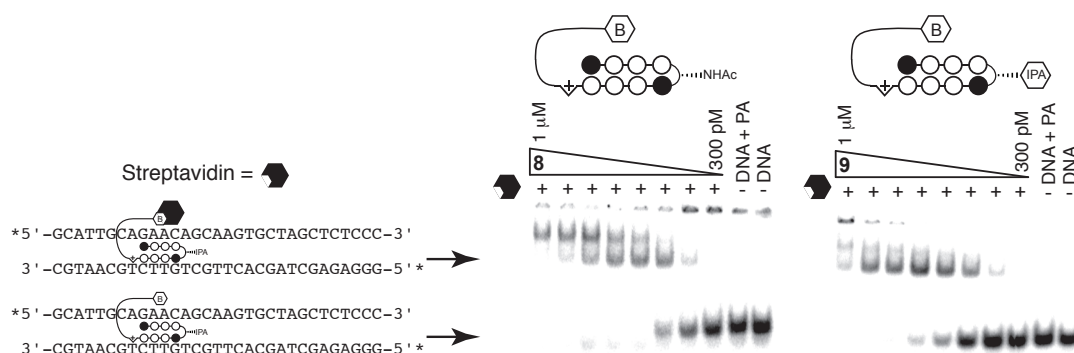
amines. All six molecules were analyzed using RT-PCR (Figure A.6). Polyamides **8** and **9** resulted in the greatest effect on gene expression. Subsequent EMSA analysis revealed that both were able to recruit streptavidin while bound to DNA (Figure A.7). Because polyamide **9** was shown to have slightly better activity in cell culture than **8**, it was identified as the lead molecule. Subsequent experiments demonstrated that polyamide **9** was also able to recruit streptavidin-bound magnetic beads while bound to radio-labeled DNA (data not shown).

DNA = 5'-TTGC <b>AGAACA</b> GCAA-3'			
Polyamides	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$	
—	60.1 ( $\pm 0.2$ )	—	
<b>A-Ac</b> 	70.1 ( $\pm 0.2$ )	10.0	
<b>C-Ac</b> 	74.9 ( $\pm 0.1$ )	14.8	
<b>4</b> 	73.3 ( $\pm 0.3$ )	13.2	
<b>5</b> 	68.6 ( $\pm 0.1$ )	8.5	
<b>6</b> 	68.0 ( $\pm 0.2$ )	7.9	
<b>7</b> 	76.4 ( $\pm 0.3$ )	16.3	
<b>8</b> 	72.4 ( $\pm 0.1$ )	12.3	
<b>9</b> 	72.1 ( $\pm 0.4$ )	12.0	

**Table A.2.** Melting temperatures of DNA/polyamide complexes. Polyamides **4-9** are compared alongside **A-Ac** and **C-Ac** in binding to a 14mer containing the PSA ARE sequence. All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses.  $\Delta T_m$  values are given as  $T_m^{(\text{DNA/polyamide})} - T_m^{(\text{DNA})}$ .



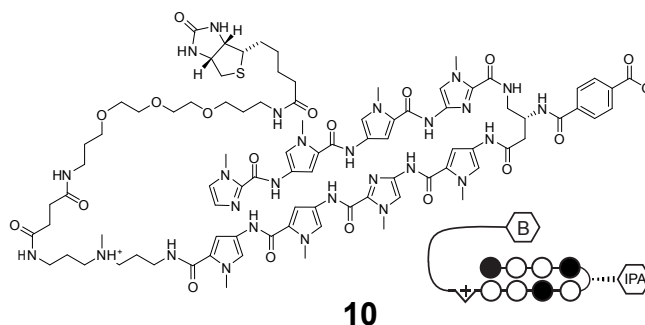
**Figure A.6.** Effects of polyamides 4-9 on PSA expression. LNCaP cells were treated with polyamides at 1  $\mu$ M and 5  $\mu$ M concentrations. The effects of these polyamides are compared with not induced (NI), induced (I), and C-Ac treatment conditions. DHT = dihydrotestosterone. Ball-and-stick models of polyamides are shown below.



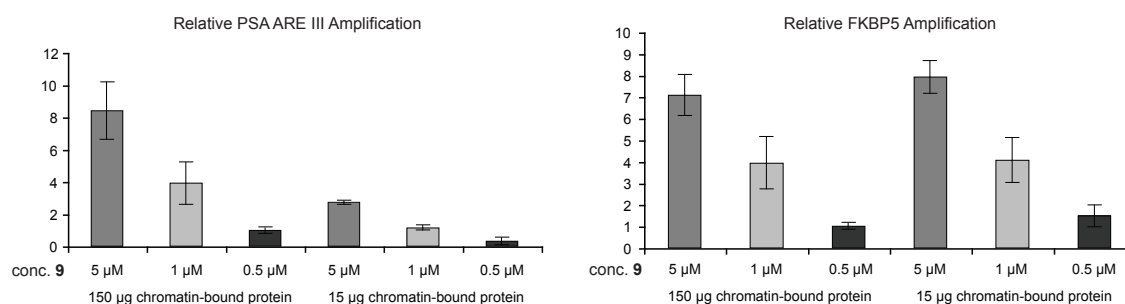
**Figure A.7.** Electrophilic mobility shift assay (EMSA) with polyamides 8 and 9. Shown is a schematic diagram of the polyamide binding the PSA ARE promoter sequence. Streptavidin is indicated by a black hexagon containing a removed wedge, and arrows indicate shift of radiolabeled DNA when streptavidin is and is not bound (left). Polyamides are titrated from 300 pM to 1  $\mu$ M, and DNA + PA (1  $\mu$ M) and DNA only controls are shown for each gel shift assay (right).

## Polyamide pull-down of DNA

Prior to pull-down experiments, polyamide **9** was synthesized on micro-molar scale in addition to polyamide **10**, which targeted 5'-WGWCGW-3' and was used as a mismatched control polyamide (Figure A.8). As a preliminary pull-down experiment, polyamide **9** was equilibrated with cross-linked, sheared chromatin (150 or 15  $\mu$ g) harvested from LNCaP cells, followed by addition of streptavidin coated magnetic beads. The solution was mixed for 24 hours at 4 °C, and beads were removed from solution and washed. The streptavidin-biotin linkage and cross-links were reversed by heating, and the resulting DNA was purified prior to analysis via RT-PCR. Interrogation of PSA ARE III and FKBP5 revealed polyamide concentration-dependant amplification of these loci (Figure A.9). Samples not treated with polyamide had undetectable amplification.

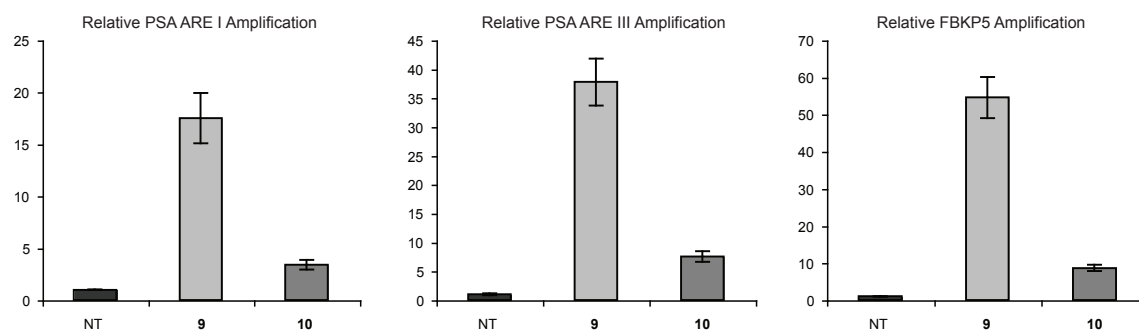


**Figure A.8.** Structure of polyamide **10**. The chemical and ball-and-stick structures for control polyamide **10** targeting 5'-WGWCGW-3' are shown. Ball-and-stick representations are as follows: an open circle is pyrrole, a closed circle is imidazole, 'IPA' within a hexagon is isophthalic acid, 'B' within a hexagon is biotin, a semi-circle represents the turn unit.

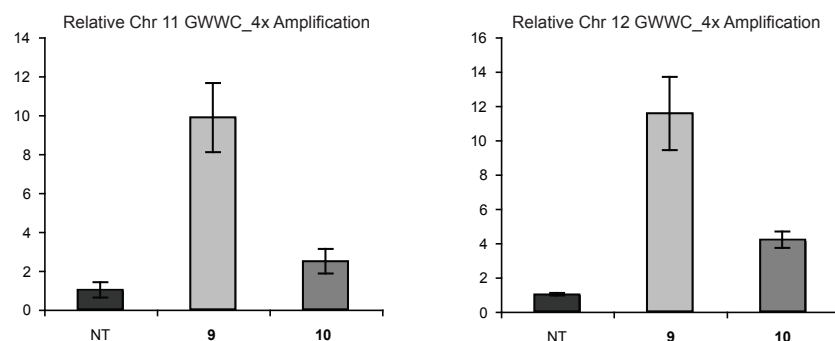


**Figure A.9.** Relative amplification levels of PSA ARE III and FKBP5 as determined by RT-PCR following polyamide incubation with sheared chromatin. Polyamide **9** was incubated with 150 µg and 15 µg chromatin-bound protein at 5 µM, 1 µM, and 0.5 µM concentrations. No amplification was detectable for non-treated samples. Samples have been normalized to the 0.5 µM/150 µg sample for each of the two loci.

Polyamides **9** and **10** were dosed in LNCaP cells at 5 µM concentration, the cells cross-linked and harvested, and the chromatin sheared by sonication as previously. Samples were split into 1 mg chromatin aliquots and streptavidin beads were added. The solutions were typically mixed for 24 hours at 4 °C, and beads were removed from solution and washed. The streptavidin-biotin linkage and cross-links were reversed by heating, and the resulting DNA was purified prior to analysis via RT-PCR. Interrogation of PSA ARE I, PSA ARE III, and FKBP5 loci showed that polyamide **9** resulted in greater amplification than untreated and **10**-treated samples (Figure A.10). Because of the promiscuity of **9** in binding to mismatched and reverse sites, no appropriate negative locus was found. We also examined regions of DNA containing four repeats of the sequence 5'-WGWWCW-3' in chromosomes 11 and 12 (Figure A.11). Relative levels of enhancement for samples treated with polyamide **9** versus untreated and **10** treated samples were not as significant. A site in chromosome 15 containing four repeats of the sequence 5'-WGWCGW-3' was also probed. Although we expected enhancement for samples treated with **10** to increase in this instance, levels were uniformly low across all treatment conditions (data not shown).



**Figure A.10.** Relative amplification levels of PSA ARE I, PSA ARE III, and FKBP5 following polyamide treatment in LNCaP cells. Polyamides **9** and **10** were dosed in cells at 5  $\mu$ M concentrations. Cells were harvested, chromatin sheared via sonication, and each treatment condition was aliquoted into 1 mg chromatin-bound protein per sample. Each sample was incubated with 2 mg streptavidin beads. Amplification is determined by RT-PCR; polyamide-treated samples were normalized to untreated control (NT).



**Figure A.11.** Relative amplification levels of loci containing 5'-WGWWCW-3' repeats. Sequences containing four repeats of the match polyamide site for **9** were interrogated. One region was located in chromosome 11 (left) and the other in chromosome 12 (right). Polyamide-treated samples were normalized to untreated control (NT).

In order to improve selectivity of pull-down by preventing non-specific interactions, bovine serum albumin (BSA) was used as a blocking agent. Prior to addition of streptavidin to samples, beads were washed with a solution of BSA in PBS. Also, a small amount of BSA solution (100  $\mu$ L) was added to each sample while mixing with streptavidin. The resulting amplification levels did not differ significantly from those of previous experiments (data not shown).

## Conclusion

To determine the genome-wide binding of polyamides in the context of chromatin *in vivo*, we have synthesized polyamide-biotin conjugates to pull down DNA-polyamide complexes using magnetic streptavidin beads. We have shown that these molecules are cell-permeable and are able to bind and recruit streptavidin *in vitro*. The chromatin from polyamide-treated cells has been harvested, sheared, and pulled-down with streptavidin beads, and the amplification of androgen responsive genes and loci containing multiple polyamide match sites has been measured using RT-PCR. It is likely that the current method of shearing DNA by sonication results in polyamide displacement and subsequent re-equilibration. Chemical or enzymatic types of shearing such as the use of micrococcal nuclease<sup>19</sup> or methidiumpropyl-EDTA (MPE)<sup>20</sup> could serve as alternative methods, however obstacles to the use of both exist. Future efforts in this arena may involve redesign of the polyamide to include a moiety capable of cross-linking to DNA so that sonication may continue to be used without polyamide displacement.

## Materials and Methods

### Polyamide Synthesis

Anhydrous *N,N*-dimethylformamide (DMF), diisopropylethylamine (DIEA), trifluoroacetic acid (TFA), and trifluoromethylsulfonic acid (TFMSA) were purchased from Sigma-Aldrich. Biotin and N-(+)-biotinyl-6-aminohexanoic acid were purchased from Sigma-Aldrich. 6-(((6-((biotinoyl)amino)hexanoyl)amino)-hexanoic acid, sulfosuccinimidylester sodium salt (biotin xx) was purchased from Invitrogen. EZ-Link TFP-PEG3-Biotin was purchased from Thermo Scientific. Bulk grade solvents were from Fisher Scientific. Preparative HPLC purification was performed on an Agilent 1200 Series instrument equipped with a Phenomenex Gemini preparative column (250 x 21.2 mm, 5  $\mu$ m) with the mobile phase consisting of a gradient of acetonitrile (MeCN) in 0.1% TFA

(aqueous). Matrix-assisted, laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on an Applied Biosystems Voyager DE-Pro spectrometer using  $\alpha$ -cyano-4-hydroxycinnamic acid as matrix.

Polyamides were generated on solid phase and in solution using previously established methods.<sup>21,22</sup> PEG-Biotin compounds were generated as follows: a solution of polyamide (8  $\mu$ mol) and Tfp-PEG3-Biotin (19.4 mg, 27.9  $\mu$ mol) in DIEA (5.6  $\mu$ L, 32.1  $\mu$ moles) and DMF (200  $\mu$ L) was allowed to stand at 23 °C for 1 h. The reaction was diluted with aqueous TFA (0.1%, 8 mL) and the crude reaction products were isolated by solid-phase extraction on a Waters C-18 Sep-Pak cartridge (10 g sorbent). The preactivated column (MeCN; followed by aqueous TFA) was loaded with the crude reaction material. The column was washed with aqueous TFA (0.1%, 50 mL) and the reaction products were eluted with MeCN (50 mL), followed by MeOH (25 mL), and concentrated *in vacuo*. The material was solubilized again with a minimal amount of MeCN, diluted with excess aqueous TFA (0.1%), frozen, and lyophilized to dryness. The benzyl carbamate (Cbz) protecting group was cleaved as previously described.<sup>17</sup> A solution of isophthalic acid (81.8 mg, 0.49 mmol) and PyBOP (133.1 mg, 0.26 mmol) in DIEA (428  $\mu$ L, 2.46 mmol) and DMF (800  $\mu$ L) was shaken at 40 °C for 10 min. Separately, the polyamide-biotin conjugate (1.87  $\mu$ L) was dissolved in DMF (500  $\mu$ L). An aliquot (200  $\mu$ L) of the PyBOP-activated isophthalic acid was added to the polyamide solution, which was allowed to stand at 23 °C for 30 min. The reaction was then diluted with a solution of MeCN in aqueous TFA (20% MeCN in 0.1% TFA, 7 mL) and purified via prep-HPLC.

Polyamide 1: MALDI-TOF  $C_{74}H_{89}N_{25}O_{14}S$  calculated  $[M+H]^+$ : 1584.73, found: 1584.60.

Polyamide 2: MALDI-TOF  $C_{80}H_{100}N_{26}O_{15}S$  calculated  $[M+H]^+$ : 1697.88, found: 1697.83.

Polyamide 3: MALDI-TOF  $C_{87}H_{112}N_{26}O_{16}S$  calculated  $[M+H]^+$ : 1810.66, found: 1810.92.

Polyamide 4: MALDI-TOF  $C_{81}H_{113}N_{26}O_{16}S$  calculated  $[M+H]^+$ : 1737.85, found: 1737.55.

Polyamide 5: MALDI-TOF  $C_{83}H_{115}N_{26}O_{17}S$  calculated  $[M+H]^+$ : 1779.86, found: 1780.82.

Polyamide **6**: MALDI-TOF  $C_{89}H_{116}N_{26}O_{19}S$  calculated  $[M+H]^+$ : 1885.10, found: 1886.78.

Polyamide **7**: MALDI-TOF  $C_{81}H_{113}N_{26}O_{16}S$  calculated  $[M+H]^+$ : 1737.85, found: 1737.50.

Polyamide **8**: MALDI-TOF  $C_{83}H_{115}N_{26}O_{17}S$  calculated  $[M+H]^+$ : 1779.86, found: 1779.81.

Polyamide **9**: MALDI-TOF  $C_{89}H_{116}N_{26}O_{19}S$  calculated  $[M+H]^+$ : 1886.10, found: 1886.64.

Polyamide **10**: MALDI-TOF  $C_{88}H_{115}N_{27}O_{19}S$  calculated  $[M+H]^+$ : 1887.09, found: 1887.01.

### **Measurement of androgen-induced mRNA**

LNCaP cells were obtained from the ATCC and maintained in RPMI 1640 media supplemented with 10% fetal bovine serum (FBS), L-glutamine, and penicillin/streptomycin. RT-PCR experiments were performed and primers used as previously described.<sup>4</sup>

### **Melting temperature analysis**

Melting temperature analysis of polyamides was performed on the indicated DNA duplexes as previously described on a Varian Cary 100 spectrophotometer equipped with a thermo-controlled cell holder, using 10 mM sodium cacodylate buffer.<sup>17</sup>

### **EMSA assay for streptavidin recruitment**

60 pmol of DNA strands (1: 5'-GCATTGCAGAACAGCAAGTGCTAGCTCTCCC-3' and 2: 5'-GGGAGAGCTAGCACTTGCTGTTCTGCAATGC-3') were annealed by heating to 95 °C for 10 min and allowing them to cool slowly to room temperature over several hours. The strands were then radio-labeled at the 5' ends using polynucleotide kinase (Roche). Labeled DNA (3,000 cpm/lane) was then incubated with polyamide at the indicated concentrations for 2 h in TAEMg buffer followed by addition of streptavidin (10 μM, Rockland) for 30 min. The total volume for each reaction was 20 μL. Running buffer (1 μL) was added to 9 μL of each reaction, and run on a 6% polyacrylamide gel and imaged using a Typhoon 8600 phosphorimager.



**Polyamide/DNA pull-down with streptavidin beads**

Approximately 90% confluent LNCaP cells grown in 175 cm<sup>2</sup> Falcon flasks were trypsinized by decanting the RPMI media, rinsing with PBS, and adding 5 mL trypsin. Following 15 min incubation, cells were diluted with 10 mL fresh RPMI, the flasks rinsed, and solution collected in a 15 mL conical tube. Cells were pelleted by centrifuging for 5 min at 25 °C, 125 RCF. The media was decanted, and 6 mL fresh media was added. The cells were pipetted up and down, and then diluted 1:10 in media for counting by hemocytometer. The concentration of the cells was adjusted to 2 million cells/mL. 24 mL of RPMI media was added to each 15-cm cell culture plate, followed by 1 mL of cells. After cells were added to plates, cells were dispersed by sliding in a vertical/horizontal motion. Cell culture plates were placed in the 37 °C incubator containing 5% CO<sub>2</sub> and were allowed to grow for 72 h. After 72 h, media from cell culture plates was decanted and replaced with 25 mL RPMI 1640 supplemented with 10% charcoal stripped FBS. Where indicated, polyamides were added in DMSO, for a total percentage of 0.1%. For each experiment five plates of each condition (polyamide- or un-treated) were generated. Plates were returned to 37 °C, and cells were allowed to grow for 64 h.

In order to fix cells, media was decanted from the plates and each was washed with ~15 mL PBS (warmed to 37 °C) prior to addition of 20 mL 1% formaldehyde in PBS. Plates were rocked for 10 min at room temperature, followed by addition of 1 mL 2.5 M glycine (0.125 M final concentration) and rocking for 5 min. The media was then decanted from plates, which were then washed with ~15 mL cold PBS (4 °C). Scraping solution (generated fresh) consists of 250 µL PMSF solution (100 mM in isopropyl alcohol) in 50 mL cold PBS. 2 mL scraping solution was added per plate, and cells were removed from plates with a razor blade and transferred via pipetting to a Falcon tube in an ice-water bath (4 °C). Cells were then pelleted by centrifugation for 5 min at 2000 RPM. Liquid was decanted, and cells were flash frozen in liquid nitrogen and stored at -80 °C.

Prior to sonication, prepared Farnham lysis buffer (FLB) and radio

immunoprecipitation assay (RIPA) buffer freshly. FLB: 5 mL 100 mM PIPES (pH 8.0), 10 mL 850 mM KCl, 5 mL 10% NP-40, 100 mL in Millipore water. Then added protease inhibitor cocktail (PIC, Roche, 2 x 50 mL tables), and sonicated to dissolve. RIPA: 10 mL 10x PBS, 10 mL 10% NP-40, 500 mg sodium deoxycholate, 1 mL 10% SDS 100 mL in Millipore water. Sonicated to dissolve. Resuspended cell pellets in 6 mL FLB each, and centrifuged cells for 5 min at 2000 RPM, 4 °C. To 10 mL of RIPA, added mini-PIC (Roche, one 10 mL tablet). Decanted supernatant and added 1 mL RIPA (containing PIC) to each pellet. Pipetted up and down to mix, and transferred to 1.7 mL eppendorf tube.

At 4 °C (in cold room) sonicated on -80 °C EtOH bath (50 mL falcon tube with EtOH stored on dry ice), 5 times 30 s at 25% amplitude. Cycled through sample tubes, switching EtOH bath with each sample, 6 times (each sample sonicated 30 times 30 s total). Between each sample sonicated with PBS for 30 s (once) to rinse. Then centrifuged sheared solutions at 14,000 RPM for 12 min at 4 °C, and transferred supernatant to a new Eppendorf tube; discarded pellet. Immediately performed Bradford assay to determine protein concentration using Cary UV-Vis and incubation with streptavidin beads.

For each sample to be measured, made 1:10 dilution in RIPA, and added 10 µL to 90 µL 0.15 M NaCl in disposable clear plastic cuvettes. For blank sample added 10 µL RIPA, 90 µL 0.15 M NaCl. For the standard curve used 2.0 mg/mL BSA, 1.5 mg/mL BSA, 1.0 mg/mL BSA, 0.75 mg/mL BSA, 0.5 mg/mL BSA (Biorad). Added 10 µL respective standard, 10 µL RIPA, and 80 µL 0.15 M NaCl. Once each cuvette has been prepared, added 1 mL 1x Bradford reagent (Biorad) to each tube in the order they are to be analyzed (including blank). Let samples sit 5-10 min at room temperature and analyzed via UV-Vis by reading at 595 nm. Measured the standards prior to samples. Generated standard curve, and interpolated concentrations of samples. Aliquoted samples such that 1 mg protein is present per tube.

For experiments where chromatin was treated with polyamides following shearing, polyamides were added at appropriate concentrations in DMSO, and allowed to equilibrate

for 12-24 h at room temperature. Then proceeded to streptavidin incubation step. Where polyamides were added in cell culture, proceeded to streptavidin incubation immediately. Some material (~ 50  $\mu$ L) was saved for input DNA sample for RT-PCR.

For experiments where binding and washing (B&W) buffer was used, prepared buffer (2x) as follows: 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, 2 M NaCl. Streptavidin beads (dynabeads, Invitrogen) were resuspended by vortexing in the original container. The appropriate amounts of beads (10 mg/mL) were pipetted into eppendorf tubes, and placed on the magnet for 1-2 min. Supernatant was removed by pipetting, and B&W buffer (twice the volume of beads originally removed) was added. The samples were vortexed, spun at 13,000 RPM for 30 s, and returned to magnet for 1-2 min. Washed 3 times. Following final removal of wash buffer, added chromatin sample to beads. If dilution is required (optimal concentration for beads is 5  $\mu$ g/ $\mu$ L), RIPA was used to obtain appropriate volume. Samples were incubated with beads for 24 h at 4 °C on end-over-end rotator.

For experiments where BSA was used as a blocking agent for streptavidin beads, prepared as follows (omitted washes with B&W buffer). Resuspended streptavidin beads by vortexing in original container. Pipetted appropriate amounts of beads (10 mg/mL) into eppendorf tubes, and placed on magnet for 1-2 min. Removed supernatant by pipetting, and added 1 mL freshly prepared BSA solution (Amresco, 5 mg/mL in PBS). Mixed on end-over-end rotator for 10 min, and place on magnet for 2 min. Removed supernatant, and repeated wash two more times. Incubated samples with beads and 100  $\mu$ L BSA for 24 h (or less where indicated) at 4 °C on end-over-end rotator. If dilution was required to reach appropriate concentration, used RIPA to obtain desired volume.

Following incubation with streptavidin beads, eppendorf tubes were centrifuged briefly and placed on magnet for 3 min. Then liquid was decanted from the tube and pellet washed 5 times with 1 mL LiCl wash buffer (10 mL Tris (pH not adjusted, stored at -20 °C), 10 mL LiCl, 10 mL 1% NP-40, 1g sodium deoxycholate, diluted to 100 mL with Millipore water). After each addition of wash buffer, placed samples on end-over-end rotator for

3 min, centrifuged, and placed on magnet for 2 min. Once final wash was completed, decanted all liquid from tube and washed beads once with 1 mL Tris-EDTA buffer (TE; 1 mL Tris-HCl pH 8.0, 0.2 mL 0.5 M EDTA pH 8.0, 998.8 mL Millipore water) by placing on end-over-end rotator for 3 min, centrifuging, and placing samples on magnet for 2 min. All liquid from tubes was decanted, and then resuspended beads in 200  $\mu$ L IP Elution buffer (1 mL 0.1 M sodium bicarbonate, 1 mL 1% SDS, 8 mL Millipore water). Eluted complexes by incubating tubes on the heated shaker at 65 °C for 1 h, vortexing every 15 min to resuspend beads. Centrifuged samples, then placed them on magnet for 2 min and removed supernatant as the IP. Added 150  $\mu$ L IP elution buffer to input DNA sample. Cross-links were reversed by transferring all samples to O-ring screw-cap microcentrifuge tubes at 65 °C for ~16 h.

To purify DNA samples for RT-PCR, extracted once with 200  $\mu$ L phenol/ $\text{CHCl}_3$ /isoamyl alcohol, vortexed for 3 s, and centrifuged at 14,000 RPM for 3 min. Transferred aqueous phase (top) to a new 1.7 mL eppendorf tube, and back extracted organic phase once with 100  $\mu$ L elution buffer, pooling both aqueous phases (back extraction is unnecessary for input DNA). Used Qiagen PCR Cleanup Kit (Qiagen) for following steps. Added three volumes of Qiagen Buffer PM, and transferred solution to Qiagen spin columns in two portions (spin for 30 s at 14,000 RPM and decanted flow-through before adding second half). Washed columns with 750  $\mu$ L Qiagen Buffer PE, spin 30 s, and discarded flow-through. Changed collection tube, and spun for 1 min at 14,000 RPM to dry column. Eluted DNA from column by adding 100  $\mu$ L Qiagen Buffer EB (warmed to 55 °C) to IP samples, centrifuging for 30 s at 14,000 RPM, and running eluate through column again for second elution. Added 200  $\mu$ L to input DNA column, and centrifuge for 30 s at 14,000 RPM.

### RT-PCR of pulled-down DNA

Each sample for RT-PCR was run in technical quadruplicate. A single PCR reaction consisted of 5  $\mu$ L template DNA and 15  $\mu$ L primer and Roche FastStart Universal SYBR Green Master (Rox). The primer/SYBR green solution was prepared as follows: 10.8  $\mu$ L of 100  $\mu$ M forward and 10.8  $\mu$ L of 100  $\mu$ M reverse primers were added to 580  $\mu$ L DEPC-treated water (USB), and the primer solution was added to SYBR green in a 1:2 ratio. A standard curve was run for each set of primers using the input DNA sample from each particular experiment performed. If multiple plates were used with a single primer pair, the standard curve was repeated on each plate. The standard curve consisted of wells containing 50 ng, 5 ng, 500 pg, and 50 pg input DNA as determined by NanoDrop UV/Vis. Dilutions were prepared in DEPC-treated water. RT-PCR was performed on an ABI 7300. Primers used were as follows:

PSA ARE I: 5'-TGCATCCAGGGTGATCTAGT-3' (forward); 5'-ACCCAGAGCTGTGGAAGG-3' (reverse)

PSA ARE III: 5'-GGATTGAAAACAGACCTACTCTGG-3' (forward); 5'-CAACAGATTGTGTTACTGTCAAGGA-3' (reverse)

FKBP5: 5'-AGCAATTTTGTGTTTGAAGAGCA-3' (forward); 5'-CTGTCAGCACATC-GAGTTCA-3' (reverse)

Negative: 5'-AAAGACAACAGTCCTGGAAACA-3' (forward); 5'-AAAAATTGCTC-ATTGGAGACC-3' (reverse)

GWWC-4x-Chr 11 (25063064-25063594): 5'-CTTTTCATGTGCAGTGTTTGG-3' (forward); 5'-CGTTTTATCAGGAAGCGTGA-3' (reverse)

GWWC-4x-Chr 12 (32538523-32539055): 5'-TTGCTGTGCAGTAATACCTAAA-3' (forward); 5'-TCCTCCATCACATCCAATCA-3' (reverse)

GWCG-4x-Chr 15 (99507897-99508423): 5'-TGAACCCCAAAACAACAAC-3' (forward); 5'-ACTCACTGGGGAGGAAGTAT-3' (reverse)

Primers for 5'-GWWC-3' and 5'-GWCG-3' repeats were designed using primer3 software (v.0.4.0)<sup>23</sup> for sites found using Python script software developed by James Puckett to search the human genome using Fuzznuc.<sup>24</sup>

## References

1. Tsai, M. J.; Omalley, B. W. *Annu. Rev. Biochem.* **1994**. *63*, 451-486.
2. Tyagi, R. K.; Lavrovsky, Y.; Ahn, S. C.; Song, C. S.; Chatterjee, B.; Roy, A. K. *Mol. Endocrinol.* **2000**. *14*, 1162-1174.
3. Roche, P. J.; Hoare, S. A.; Parker, M. G. *Mol. Endocrinol.* **1992**. *6*, 2229-2235.
4. Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2007**. *104*, 10418-10423.
5. Trauger, J. W.; Dervan, P. B. *Meth. Enzymol.* **2001**. *340*, 450-466.
6. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B.; Puckett, J. W.; Wang, C. C.; Dervan, P. B. *Tetrahedron.* **2007**. *63*, 6146-6151.
7. Warren, C. L.; Kratochvil, N. C. S.; Hauschild, K. E.; Foister, S.; Brezinski, M. L.; Dervan, P. B.; Phillips, G. N.; Ansari, A. Z. *Proc. Natl. Acad. Sci. U. S. A.* **2006**. *103*, 867-872.
8. Puckett, J. W.; Muzikar, K. A.; Tietjen, J.; Warren, C. L.; Ansari, A. Z.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**. *129*, 12310-12319.
9. Aparicio, O. M.; Geisberg, J. V.; Sekinger, E. A.; Yang, A.; Moqtaderi, Z.; Struhl, K. *Curr. Protoc. Mol. Biol.* **2005**. *69*, 21.23.21-21.23.17.
10. Fan, X. C.; Lamarre-Vincent, N.; Wang, Q.; Struhl, K. *Nucleic Acids Res.* **2008**. *36*, e125.
11. Buck, M. J.; Lieb, J. D. *Genomics.* **2004**. *83*, 349-360.
12. Johnson, D. S.; Mortazavi, A.; Myers, R. M.; Wold, B. *Science.* **2007**. *316*, 1497-1502.

13. Wilchek, M.; Bayer, E. A. *Anal. Biochem.* **1988.** *171*, 1-32.
14. de Boer, E.; Rodriguez, P.; Bonte, E.; Krijgsveld, J.; Katsantoni, E.; Heck, A.; Grosveld, F.; Strouboulis, J. *Proc. Natl. Acad. Sci. U. S. A.* **2003.** *100*, 7480-7485.
15. Dejardin, J.; Kingston, R. E. *Cell.* **2009.** *136*, 175-186.
16. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007.** *35*, 363-370.
17. Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B. *J. Am. Chem. Soc.* **2008.** *130*, 6859-6866.
18. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B. *Angew. Chem.-Int. Edit.* **2007.** *46*, 7956-7959.
19. Horz, W.; Altenburger, W. *Nucleic Acids Res.* **1981.** *9*, 2643-2658.
20. Hertzberg, R. P.; Dervan, P. B. *Biochemistry.* **1984.** *23*, 3934-3945.
21. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B. *Bioorg. Med. Chem.* **2002.** *10*, 2767-2774.
22. Chenoweth, D. M.; Harki, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **2009.** *131*, 7175-7181.
23. Rozen, S.; Skaletsky, H. J. **2000.** *Bioinformatics Methods and Protocols: Methods in Molecular Biology*, 365-386.
24. Rice, P.; Longden, I.; Bleasby, A. *Trends Genet.* **2000.** *16*, 276-277.